MODELLING SPATIAL TURNOVER PATTERNS IN TREE SPECIES DIVERSITY FOR CONSERVATION PLANNING IN THE MESOAMERICAN REGION

IVIS JULIETA CHAN AGUILAR

Thesis submitted for the degree of Doctor of Philosophy

Bournemouth University
In collaboration with
PBL Netherlands Environmental Assessment Agency

JULY 2015

Abstract

The planet's biodiversity is increasingly threatened by the effects of rapid environmental change. Biodiversity loss led by expanding anthropogenic activity, combined with the effects of changing climatic regimes, pose a threat to natural communities and the ecosystem services that they underpin. Understanding how communities of species vary across space and the factors that drive such variation is essential to predicting the impacts of human and climate-induced changes on biodiversity and ecosystem services. This knowledge can be used to plan conservation areas that effectively represent the underlying ecological processes that maintain biodiversity and ensure the provision of ecosystem services crucial to human well-being.

Beta diversity has been variably defined since the concept was first introduced as a measure of variation in species assemblages. This has given rise to the development of different measures of beta diversity, and to varying conclusions regarding the variation of species composition across space. In this thesis, I aimed to investigate the spatial structuring of tree species turnover patterns (beta diversity) and the potential drivers of these patterns, within the highly environmentally heterogeneous Mesoamerican region. In addition, the effect of habitat loss on tree diversity up to the year 2000 was estimated, and the representativeness of the protected area system was assessed in an effort to highlight potential areas of conservation concern across the region. The efficacy of ecoregions in representing turnover patterns was also assessed.

The spatial structuring of areas of low similarity (high beta diversity) was clustered in the north and south of the study region, as well as within the Central American mountainous regions. Areas of low similarity mostly fell within dry, pine-oak, and montane ecoregions. Moist forests were indicated as having high similarity. The congruence of the modelled pattern of beta diversity in trees with patterns of high beta diversity in other taxa provides insights into areas of potential conservation efforts. Habitat loss in the past decade was high in the southern areas of the isthmus indicating that countries like Nicaragua, Costa Rica, and Panama had lost high proportions of their original habitat and species. However, habitat loss was intensive throughout the region, and areas of high beta diversity were found to be among the least represented in the current protected areas system. Classification based on beta diversity broadly resembled the WWF ecoregional classification for the region. This suggests that models based on broad climatic variables can help elucidate beta diversity patterns at broad spatial scales, although conservation planning also requires robust information at finer scales.

List of Contents

List of Figures

List of Tables

Acknowledgements

The old adage says it takes a village to raise a child. From my experience throughout this process, I would venture to say it takes a world to raise a scientist. All the people I have met in my life have contributed in some way to the culmination of this arduous and rewarding process. Such a list would not fit in these pages, so I limit myself to thanking only some of those people here.

First of all, I thank and acknowledge my Bournemouth University supervisors: Professor Adrian Newton and Dr. Duncan Golicher. Without their generous patience, expert guidance and moral support, I would not have been able to continue this journey. They have taught me invaluable lessons about the challenges and joys of actually "doing" science. Although we may have disagreed many times and their patience may have run thin, I thank them for keeping their confidence in me. I also thank my supervisor, Dr. Rob Alkemade, and collaborator, Dr. Richard Verweij at PBL Netherlands Environmental Assessment Agency. My time in the Netherlands working with them is one of the happiest I take from this journey. I also thank PBL Netherlands Environmental Assessment Agency for partly funding this work and allowing me to pursue this research. I also take this opportunity to thank the staff at BU AppSci and the Graduate School, especially Louise Pearson, Professor Holger Schutkowski and Fiona Knight whose guidance and support helped me immensely in navigating the world of academia.

I give special, massive thanks to those I warmly consider the Australian posse: Dan Rosauer, Shawn Laffan, Glenn Manion, and Tom Harwood. My colleague, friend and honorary supervisor, Dan, has been a source of inspiration and voice of reason for me throughout this arduous process. I thank him for being so generous with his expertise and time, and for always being positive and supportive when I constantly complained about things. I am lucky that he let me drink from his well of eternal optimism. And I am grateful that he let me spend time with his lovely family, especially Vida and little Oscar. I will always remember that time when we went kangaroo spotting in Canberra. Special thanks to Shawn, who not only shared with me his expertise on his software baby, BIODIVERSE, but also always offered good advice and moral support. I will retain fond memories of visiting the headlands in Sydney with his family, especially chatty little Maya and Francis.

I am also especially grateful to Luis Cayuela and the BIOTREE network for facilitating access to data for this research project and for invaluable advice. I owe thanks also to the Jetz lab at Yale University, especially Professor Walter Jetz for giving me the opportunity to pursue some of my research at his lab, and work with his colleagues. I am grateful to have met many talented scientists at his lab, especially; Jonathan Bellmaker, Morgane Barbet-Massin, and Petr Keil, whose discussions during lunch time gave me lots to think about and helped me refine some of my research questions. I also wish to thank Professor William Kunin and Tim Baker at the University of Leeds for starting me off on this beta diversity strand of research.

I am forever grateful to all the friends who have always been there to support me and make me laugh, no matter how dark the days seemed: Heather, Carina, Zoe, Norman and Hannah, Kat Evans, Katie, Natalia, Carolina, Nicola, Daraporn, Matt, Justine, Quyen, Kathryn, Laura, Eva, Grace, Citlalli, Isabel, Ari, Daniel, Sergio, Alexandra, Zenek, Greg, Clare, Alex L, Marie Christine, Kate and Adam, and Carrie and Paul Hughes. I am especially thankful to all those friends who have made the last few months and days of this journey especially pleasant and joyous (because it could have been the opposite!): Emily, Hannah, Jenny, Arjun, Paul and Katherine B. I am grateful also to some special friends who have shared their experience, strength, and hope with me in the last few years. I thank them for their unconditional support: Sharon Hartwell, Janet, Jean, Julee, Jo, Mary, Sally, David, Ian, Jacqui, Chris and Lesley.

I thank my Florida family, Jack and Carmen, who are my safe place when life gets a little too stormy. I am so fortunate that they have adopted me as part of their family. I am greatly thankful to my special friend, Alan, for his massive support throughout this process. His curiosity, child-like sense of wonder and sense of humour have kept me going many a day when I wanted to give up. Some of my happiest moments have been exploring the UK with you!

I thank my family at home for their support and for telling me they miss me all the time I have been away. I am grateful to my family and friends who have inspired me but have unfortunately left this dimension: Bueli Tami & Bueli Pol, George Singh, Aura Luz Alcoser. I especially remember my cousin, Fer, who passed away too young. This thesis is dedicated to you, for our childhood.

Finally, I thank my mom for her unconditional love and support always. She is my rock and I would not be the woman I am if it wasn't for her unwavering belief in me.

Author's Declaration

I confirm the work presented in this thesis is my own work. The R scripts used throughout the thesis were derived from the open-source R Project for Statistical Computing (http://www.r-project.org/) and modified in collaboration with my supervisor, Duncan Golicher. Additional software, data sources and materials used have been fully acknowledged throughout the thesis.

1 CHAPTER 1

Introduction

1.1 Biodiversity & the Three Cs: Climate Change, Carbon Storage & Conservation Planning

The survival and well being of human society has always been closely linked to nature. People depend on nature through its provision of ecosystem services such as clean water and air, fertile soils for food production, or recreation as well as less visible services such as flood control and climate regulation (Rodríguez et al. 2006). Our continuing manipulation of the environment has had significant effects on nature's ability to continue providing such services. Anthropogenic activity is responsible for widespread environmental degradation, and is also considered to be a major driver of climate change owing to increasing CO_2 emissions. The impacts of climate change on biodiversity are well documented, based on studies of different taxa conducted in recent years. Species in ecosystems particularly sensitive to changes in regional climate, such as high-altitude and high-latitude areas, have shown significant responses to climate change through changes in their breeding seasons, distribution and migration patterns, and shifts in their ranges (Gitay et al. 2002). Such changes have been observed in many taxa including amphibians, insects, birds, mammals and plants.

Tropical forests, which constitute the most biodiverse and carbon rich ecosystem on earth, are particularly threatened by the effects of both human land-use conversion activities and climate change (Strassburg et al. 2010). Climate-induced changes in dry season or fire regimes, for example, can lead to biodiversity loss and reductions in biomass and carbon stocks (Midgley et al. 2010). The Mesoamerican region is expected to undergo varied changes under climate change including up to 40% decrease in precipitation, longer dry seasons and higher hurricane intensities (Galindo et al. 2010). Neotropical forests within the region are thought to play an important role in balancing atmospheric carbon concentration by acting as sinks of CO_2, as well as acting as sources of CO_2 after deforestation (Malhi and Grace 2000). Their contribution to mitigating climate change is reduced by deforestation, which is a major contributor to global carbon emissions, second only to the combustion of fossil fuels (Ebeling and Yasué 2008). Efforts to protect forests in tropical countries through carbon payment schemes such as REDD+ (Reducing Emissions from Deforestation and Degradation) are therefore seen as a major step in reducing global carbon emissions (Phelps et al. 2010). The magnitude of potential funding flowing from such carbon payment schemes could rival profits from forestry and therefore provide a major incentive towards forest conservation (Miles & Kapos 2008). Such carbon payment schemes are also attractive because of their potential to provide additional ecosystem benefits such as biodiversity conservation and watershed protection (Dickson et al. 2009).

It seems intuitive that protecting forests to reduce carbon emissions will also conserve biodiversity. Recent research, however, suggests that co-benefits such as biodiversity conservation will only ensue if carbon-based projects explicitly include biodiversity benefits as a factor in design processes (Grainger et al. 2009; Harvey et al. 2010). Tropical land use conversion (deforestation) is a major cause of terrestrial biodiversity loss, and is not just a contributor to climate change (Ebeling and Yasué 2008; Steffan-Dewenter et al. 2007). It is therefore important to reconcile conservation actions aimed at conserving

forests to reduce carbon emissions with conservation efforts aimed at reducing biodiversity loss. In carbon-based schemes such as REDD+, for example, decisions of which areas to conserve may primarily be driven towards conserving areas that will provide high returns in terms of reducing carbon emissions and where implementation costs are low (Ebeling and Yasué 2008; Venter et al. 2009). Such areas may not necessarily coincide with areas of high conservation value (harbouring high levels of endemic and/or threatened species). Conversely, areas of high conservation value may assume higher implementation costs for low carbon returns (Ebeling and Yasué 2008).

In tackling this new issue in conservation planning, a cautious approach must be applied to ensure that conserving for carbon does not imply a loss to biodiversity conservation, but in fact maximizes biodiversity co-benefits. Examining the ecological trade-offs between biodiversity and carbon storage presents an opportunity for ecologists to contribute vital knowledge to facilitate informed decision-making under this new framework in forest conservation (Baker et al. 2010). Driven by the observation that mitigating biodiversity loss has widely been seen only as a side-benefit in climate change mitigation, research attention has increasingly focused on substantiating the biological links between diversity and carbon storage (Díaz et al. 2009). Recent research highlights empirical evidence for the influence of biodiversity on carbon sequestration. For example, dominant plant species have a significant influence on the size and turnover rate of biomass carbon (Díaz et al. 2009). There is also sufficient evidence to suggest that species richness, identity, relative abundance and spatial arrangement may all have an impact on carbon sequestration capacity (Díaz et al. 2009). Such observations underline the ecological basis for integrating biodiversity protection with carbon-based forest conservation initiatives.

Clearly, the issue for decision-makers lies in balancing carbon interests with conservation to mitigate biodiversity loss, in a cost-effective manner. Historically, protected areas have been established for a variety of reasons from protecting iconic species (e.g. jaguars) or fragile ecosystems (e.g. wetlands) to meeting specific global conservation targets (e.g. Important Bird Areas) (Hole et al. 2009). This of course leads to different areas being prioritized for different reasons. More recent systematic approaches have applied various prioritization schemes including addressing specific threats, mapping separate land classes to be used as surrogates for representing biodiversity, and combinations of biodiversity representation and degree of threat (Margules & Pressey 2000; Ferrier 2002; Wilson et al. 2007). Regardless of approach, most protected area networks are based on static views representing present patterns in species composition and distribution (Hole et al. 2009). However, the recognition that climate change poses a great threat to endemic species with restricted climatic ranges (Gasner et al. 2010), and that such effects might become more severe in the future (Golicher et al. 2012), makes it imperative to understand how present patterns of species diversity may be altered under climate change. From a conservation planning perspective, the implication is that present conservation regions may, in the future, no longer conserve certain genetic or species diversity, or ecosystems (Gitay et al. 2002). This in turn necessitates a planning approach aimed at effectively representing species diversity patterns and preserving the persistence of processes that maintain them (Pressey et al. 2007).

Conservation areas should thus be designed to ensure long-term persistence of the ecological elements and processes that maintain biodiversity (Fairbanks et al. 2001) and give rise to provision of ecosystem

services. It is generally accepted that increasing the number of reserves would capture more diversity (Wiersma and Urban 2005); however, this begs the question of whether additional reserves would be capturing different suites of species or if they would just encompass more of the same species. The establishment of representative conservation areas therefore depends on the degree of species compositional heterogeneity, or beta diversity, within the region of interest (Wu et al. 2010). The need to incorporate environmental heterogeneity in conservation planning has been previously recognized (Cowling & Pressey 2001; Londoño-Murcia et al. 2010), but the opportunity to incorporate beta diversity as a measure towards this is only recently being explored. The concept of beta diversity is receiving growing research attention, and there is a growing body of literature exploring the application of beta diversity in the context of conservation and climate change planning (Allnutt et al. 2008; Dunlop et al. 2012; Fairbanks et al. 2001; Faith et al. 2003, 2004; Ferrier and Drielsma 2010; Ferrier 2002; Ferrier et al. 2002, 2004, 2007a; Fitzpatrick et al. 2013; Kraft et al. 2011; Leathwick et al. 2010; Mokany et al. 2011; Steinitz et al. 2005; Wiersma and Urban 2005; Williams et al. 2010; Wu et al. 2010).

Incorporating beta diversity into conservation planning is a logical next step since beta diversity not only highlights areas of unique species composition, but could also be used as a surrogate for ecological processes since it reflects diversification along specific environmental gradients (Fairbanks et al. 2001). This study will therefore explore the application of beta diversity in conservation planning, by focusing on measuring beta diversity as a necessary component in differentiating areas for conservation purposes.

1.2 Defining Beta Diversity

Conservation priorities are often determined by quantifying species diversity at some spatial scale. Species diversity has been measured in several ways including endemism, genetic diversity and species richness (Crisp et al. 2002; Laffan & Crisp 2003). Species richness can itself be distinguished into three components: alpha, beta and gamma diversity (Koleff & Gaston 2002). Alpha diversity represents local species richness and gamma diversity represents regional species richness (Koleff & Gaston 2002). Beta diversity, or spatial turnover, is the variation in species composition among localities (Koleff & Gaston 2002). The scale at which these components of species diversity are measured determines which areas are characterized with high diversity in relation to other areas, and therefore holds important implications for the targeting of conservation efforts.

In his seminal papers (Whittaker 1960, 1972), Whittaker defines beta diversity in essentially two ways (i) as the ratio between regional (gamma) diversity to local (alpha) diversity, and (ii) the variation in species composition along habitat gradients, usually measured through pair-wise comparison of focal areas. These definitions followed from the concept of niche theory which argues that species occupy different positions along habitat gradients based on differences in their utilization of resources (Whittaker 1972).

Whittaker also defined a multiplicative approach to the relationship between alpha and gamma diversity, framing gamma diversity as the product of local community diversity and the degree of beta diversity between communities (Whittaker 1972). Others have applied an additive approach, defining gamma diversity as the sum of local diversity and beta diversity (among-community diversity) (Veech et al.

2002; Baselga 2010). This essentially equates beta diversity to the average amount of regional diversity not found within the focal sample (Veech et al. 2002), also called the effective number of species absent from the average community (Jost et al. 2010). This is purported to be an easier interpretation of beta diversity that allows direct comparison between alpha and beta diversities since both are averages (Veech et al. 2002). Attempts to clarify the two definitions of beta diversity first offered by Whittaker have distinguished between the first definition as 'proportional' diversity and the second as 'differentiation' diversity (Anderson et al. 2011; Jurasinski et al. 2009). However, it is now argued that differentiation diversity (with its many pair-wise similarity indices) is not a different concept but is really encapsulated within the multiplicative approach of proportional diversity, for the case of comparing any two samples (Baselga 2010).

Further reviews of the concept of beta diversity examine the varying definitions that have been applied since the concept was first introduced, and summarize the ways in which these different views have arisen. These include defining diversity itself in new ways, defining the alpha component in different ways, defining the relationship between gamma and alpha in a new way, and forming a totally new construct of beta diversity with no relationship to alpha and gamma (Tuomisto 2010). Tuomisto (2010) proposes a conceptualization of beta diversity based on defining *true species diversity* itself as the number of 'effective' species that a dataset represents, having taken into account the mean proportional abundance of the actual species (Tuomisto 2010). This is based on the gamma classification of entities (e.g. individual trees) into types (species), and the omega classification of these same entities (individuals) into the sampling units in which they were observed (Tuomisto 2010). In this sense, then, *true beta* diversity quantifies the number of "compositionally distinct virtual sampling units that have the same species diversity as the actual sampling units do on average" (Tuomisto 2010). In other words, *true beta* diversity represents the classification of effective species into distinct units which share no effective species with each other.

Such different perspectives on the concept of beta diversity have yielded different approaches for investigating compositional variation in species assemblages, and consequently different measures of beta diversity. It is not surprising, then, that studies of beta diversity patterns have demonstrated varied results. Results suggest that emergent beta diversity patterns ultimately depend on characteristics of the study organism (e.g. dispersal mechanism) (McKnight et al. 2007; Sesnie et al. 2009), the scale at which the study is conducted (Gering & Crist 2002; González-Espinosa et al. 2004; Veech & Crist 2007), and the individual and interactive influence of underlying processes (e.g. distance or environmental gradients) that may be driving such patterns (Jankowski et al. 2009; Melo et al. 2009; Steinitz et al. 2006).

Clearly, any conclusions drawn from analysis of beta diversity patterns on a dataset depends on the research question being asked, including the concept of beta diversity underlying it and the choice of how to measure this beta diversity.

1.3 Measuring beta diversity

Koleff et al. (2003) provides an extensive review of many of the beta diversity measures that have been applied to the study of beta diversity patterns. Koleff et al. (2003) provides a clear conceptualization of beta diversity measures or indices by re-expressing them according to "matching components" (**Figure 1.1**) where 'a' equals the number of shared species between the two samples being compared, 'c' equals the number of species found only in the focal sample (species loss) and 'b' equals the number of species found only in the neighbouring sample (species gain). The authors further identify classes of indices depending on which of the matching components is emphasized (**Table 1.1**).

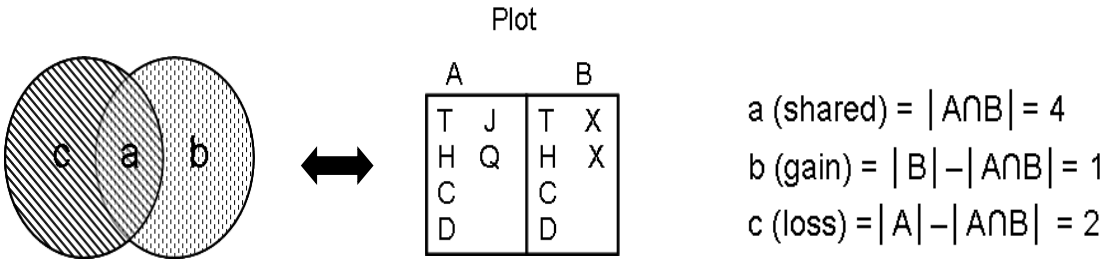


Figure 1.1 The matching components used to calculate similarity/dissimilarity indices (adapted from Koleff et al. 2003) expressed in set notation. Taking Plot A as the focal sample, component 'a' is the number of elements in the intersection of plot A & B. Component 'b' is the number of elements in plot B minus component 'a'; component 'c' is the number of elements in plot A minus component 'a'. Note differing species richness; b < c (species gain < species loss).

Table 1.1 A sample of various beta diversity indices. Adapted from Koleff et al. (2003) and Baselga (2010). Formulae are re-expressed in terms of matching components as per Koleff et al. (2003), except for M_{sim} (Baselga et al. 2007). Broad-sense indices reflect differences in species richness, whereas narrow-sense indices reflect species turnover/replacement.

Measure	**Formula**	**Explanation**
B_w (Whittaker's beta)	$B_w = \frac{a+b+c}{(2a+b+c)/2}$	Broad-sense (Koleff et al. 2003).
B_j (Jaccard)	$B_j = \frac{a}{a+b+c}$	Broad-sense (Koleff et al. 2003).
B_{sim} (Simpson)	$B_{sim} = \frac{\min(b,c)}{a+\min(b,c)}$	Narrow-sense; values between 0-1; high values=high dissimilarity=high beta diversity; similarity independent of richness differences; measures beta diversity owing to species replacement/turnover (Baselga 2010).
B_{gl} (Lennon)	$B_{gl} = \frac{2[\max(b,c)-\min(b,c)]}{2a+\max(b,c)+\min(b,c)}$	Developed to discern differences in beta diversity owing to differences in species richness; not applicable to parsing nestedness vs species turnover (replacement) patterns as it includes all richness differences not just those attributable to nestedness (Baselga 2010).
B_{sor} (Sorensen)	(i) $B_{sor} = \frac{2a}{2a+b+c}$ (Koleff et al. 2003) (ii) $B_{sor} = \frac{b+c}{2a+b+c}$ (Baselga 2010)	Broad-sense; values between 0-1. Low values=low similarity=high beta diversity. Different numeric values are produced due to differential weight given to (i) shared species vs (ii) gained or lost species. Formula (ii) also known as B_t, reformulation of B_w (Koleff et al. 2003). B_{sor} is a similarity measure; the dissimilarity would be 1-B_{sor}.
B_{nes} (nestedness)	$B_{nes} = \frac{\max(b,c)-\min(b,c)}{2a+\min(b,c)+\max(b,c)} \times \frac{a}{a+\min(b,c)}$	Developed to measure beta diversity due to nestedness (Baselga 2010).
M_{sim}	$M_{sim} = \frac{\sum_i S_i - S_T}{\left[\sum_{i<j} \min(b_{ij}, b_{ji})\right] + \left[\sum_i S_i - S_T\right]}$	Multiple-site similarity index (Baselga et al. 2007), where S_i = total number of species in site i and S_T = total number of species in all sites considered together; b_{ij} and b_{ji} = number of species exclusive to sites i & j respectively, when compared by pairs.

Measures or indices of continuity, or broad-sense indices, give emphasis to the number of species shared between the two samples (component 'a') reflecting beta diversity as the relationship between alpha and gamma diversity, rather than as a measure of difference in species composition (Koleff et al. 2003). The value of such indices will vary at high or low extremes depending on whether little or many species are shared and according to the difference in richness between samples (**Figure 1.2, a & b**) (Koleff et al. 2003). Consequently, these indices tend to reflect differences in species richness rather than differences in species composition (Koleff et al. 2003). On the other hand, indices of gain and loss (narrow-sense indices) depend on component 'a' but also take into account the magnitude of species gains and losses made in moving from the focal sample to the neighbouring sample (**Figure 1.2, c & d**). Hence, they measure species turnover and are less reflective of differences in species richness (Koleff et al. 2003). For this reason, Koleff et al. (2003) advocates the use of such narrow-sense indices (e.g.) Simpson dissimilarity index.

a)

c=100	a=10	b=200

$$B_{sor} = \frac{2a}{2a+b+c} = \frac{2\times10}{(2\times10)+200+100} = 0.063$$

$$B_{sim} = \frac{\min(b,c)}{a+\min(b,c)} = \frac{100}{100+10} = 0.91$$

b)

c=100	a=10	b=1000

$$B_{sor} = \frac{2a}{2a+b+c} = \frac{2\times10}{(2\times10)+1000+100} = 0.018$$

$$B_{sim} = \frac{\min(b,c)}{a+\min(b,c)} = \frac{100}{100+10} = 0.91$$

c)

T	T H C D J

$$B_{sor} = \frac{2a}{2a+b+c} = \frac{2\times1}{(2\times1)+0+4} = 0.33$$

$$B_{sim} = \frac{\min(b,c)}{a+\min(b,c)} = \frac{0}{0+1} = 0$$

d)

T H C	T D J

$$B_{sor} = \frac{2a}{2a+b+c} = \frac{2\times1}{(2\times1)+2+2} = 0.33$$

$$B_{sim} = \frac{\min(b,c)}{a+\min(b,c)} = \frac{2}{2+1} = 0.66$$

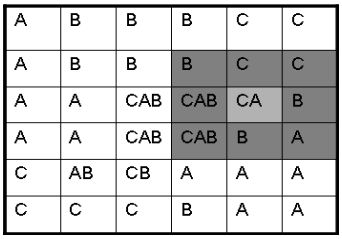
Figure 1.2 Examples of the difference between B_{sor} and B_{sim} indices. Boxes represent grid cells containing species. The letter 'a' is the number of shared species, letters 'b' and 'c' are the number of unshared species in respective cells. In c) and d) letters inside grid cells represent species. In a) and b), B_{sor} reflects the magnitude of the difference in species richness between grid cells, whereas B_{sim} reflects only the turnover or replacement of species between grid cells. In c) and d) B_{sor} does not distinguish the nestedness situation in c) and shows the same numerical value for both situations. In contrast, B_{sim} does distinguish between both scenarios and assigns a zero value to the nestedness situation.

The Simpson dissimilarity (B_{sim}) index (a narrow-sense index) indicates that in **Figure 1.2 a-b** the plots are equally very dissimilar (1 is total dissimilarity). The index value does not vary with the difference in plot species richness like the Sorensen similarity (B_{sor}) index (a broad-sense index). The Sorensen metric produces different numeric values depending on the species richness between plots, i.e., it is sensitive to species richness differences. In **Figure 1.2 c-d**, the Sorensen index treats both cases as equal; the numeric value of the Sorensen index is the same in both cases, indicating no difference between them. Component 'a' is the same in both cases but the magnitude of gains and losses is not accounted for by this measure. In contrast, the Simpson index does take into account these gains and losses and reflects this in different numeric values produced for the two cases. In this instance, the Simpson index indicates there is no species turnover in case c), since the species pool of one cell is a subset of the species pool of the neighbouring cell.

Beta diversity patterns can reflect nestedness (**Figure 1.2, c**) of species assemblages or species turnover (Baselga 2010). Nestedness occurs when sites of lower species richness are subsets of sites of greater species richness, whereas species turnover refers to replacement of some species by others among sites (Baselga 2010). A given community may reflect nestedness, turnover or any combination of these (**Figure 1.3**). Since these differing patterns may reflect different underlying historical processes such as

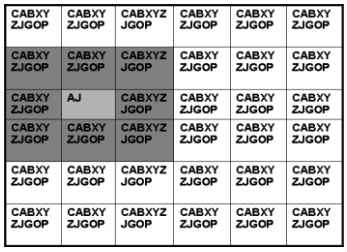
the presence of refugia or speciation, it is advisable to differentiate between patterns of nestedness and species turnover (Baselga 2010).

a)



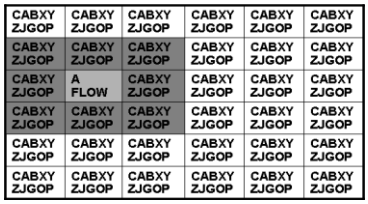
Low beta diversity (both β_{sim} and β_{sor}); nestedness of biotas: the focal cell is a subset of the neighbour cell species pool but they are not much different in species richness. There is no species replacement here.

b)



β_{sim} reaches minimum value of zero (total similarity); nestedness of biotas: the focal cell is much less species rich and is a subset of the neighbour cell species pool. Beta diversity represents species loss from the focal cell to the neighbour cell.

c)



High beta diversity (both β_{sim} and β_{sor}); nestedness and species turnover. The focal cell contains species from the neighbour species pool but also new species.

d)

High beta diversity (both β_{sim} and β_{sor}); species turnover/replacement as defined by Baselga (2010) not nestedness since the focal cell represents an entirely different species pool.

Figure 1.3 Beta diversity patterns due to nestedness and/or species turnover (replacement).

Patterns of beta diversity due to nestedness can be discerned using the pairwise nestedness dissimilarity measure (B_{nes}) proposed for this purpose (**Table 1.1**), where $B_{sor}=B_{sim}+B_{nes}$ therefore $B_{nes}=B_{sor}-B_{sim}$**.** This is based on the proven ability of B_{sim} to differentiate turnover from nestedness, and thus B_{sor} is a measure that reflects the combined effects of turnover and nestedness (Baselga 2010). The latter author explains that B_{nes} is a "measure of the dissimilarity of communities due to the effect of nestedness". It is not an absolute measure of nestedness. That is, it cannot be used to measure nestedness of a pair of samples as opposed to the turnover in the same pair of samples. Rather, it provides a way to discern dissimilarity between communities of differing species richness, when beta diversity patterns are caused by nestedness (**Figure 1.4**).

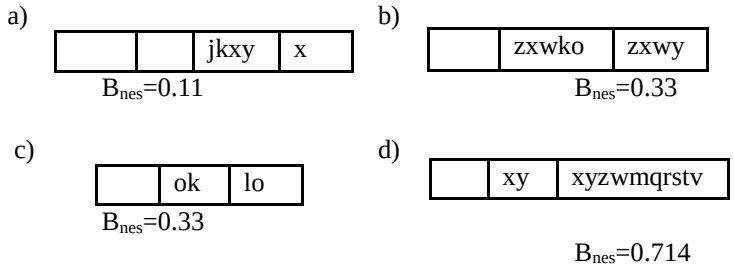


Figure 1.4 Beta diversity due to nestedness (B_{nes}). Boxes represent grid cells containing species (letters). Note that cases b) & c) have equal values of B_{nes} because these two situations have the same proportion of species shared and not shared.

B_{nes} increases from a) to d) in **Figure 1.4**, owing to increasing differences in the number of species between plots. Therefore, **Figure 1.4d** shows highest dissimilarity as it has the greatest difference in species richness between plots. Here we can see that nestedness does not incorporate species replacement and will only give rise to high beta diversity if the biota of one site is significantly richer than the other (as in **Figure 1.3b**).

High beta diversity can mean that many different species occupy the same space. If the 'range' of a particular species is defined as the number of cells/samples in which that species occurs within a particular landscape, then the spatial arrangement of species ranges within that space can be different (**Figure 1.3**). Consequently, different configurations of species assemblages and thus beta diversity patterns can arise. Overlap of species ranges, that is, nestedness of biotas will reflect low beta diversity (as measured by β_{sim}) regardless of the difference in species richness between the two samples being compared (**Figure 1.3, a & b**). On the other hand, high beta diversity is consistently shown in cases where species replacement occurs. Species turnover is displayed when the ranges of some species are entirely contained within the ranges of others (**Figure 1.3**, **c & d**). At a broad scale, this may reflect patterns of endemism (Laffan & Crisp 2003; Baselga 2010). At a more local scale, this does not necessarily have to be tied to endemism since such patterns can arise at ecotones—transition zones between vegetation assemblages (**Figure 1.5**).

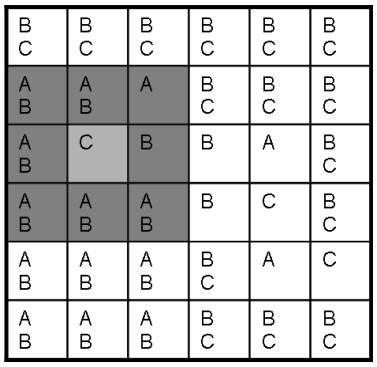


Figure 1.5 Vegetation transition zone (ecotone). B_{sim} indicates total dissimilarity (reflects high beta diversity) between the focal cell (light gray) and its neighbouring cells (darker gray).

Partitioning beta-diversity into nestedness and replacement components can be used to identify whether conservation areas represent species at the core of their ranges or whether they encompass only transition areas (**Figure 1.6**). Transition areas can represent areas of high *alpha* diversity due to the convergence of species ranges (nestedness) or they can show species replacement if the transition area represents one type of vegetation giving way to another type (ecotone) e.g. broadleaf forest to savannah. Partitioning beta-diversity into nestedness and turnover components allows us to identify which of these two scenarios is occurring. Coupling this with visualization of species ranges (where species occur) allows us to see whether areas that may be proposed for conservation are actually encompassing species at the core or edges of their ranges, and how representation of species can be maximized.

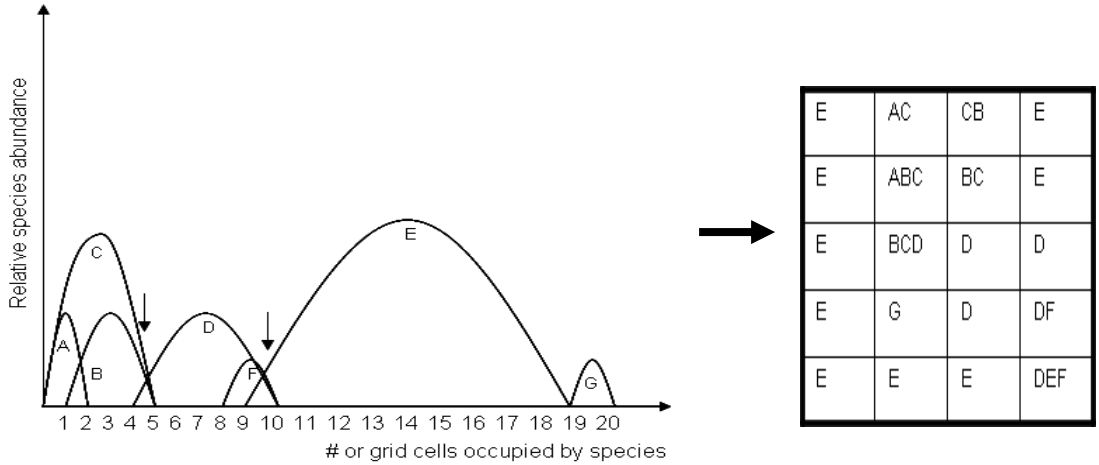


Figure 1.6 In the graph, the areas pointed by the two arrows represent nestedness, whereas the area in G represents species replacement (species range does not overlap with any others). Note that the spatial configuration of the species in the grid cells can be different. (Graph modified from Gaston et al. 2001).

Since areas showing high beta diversity essentially represent transition areas of different composition (once nestedness is accounted for), identifying beta diversity patterns as a result of species turnover can assist in making better judgements about conservation areas. This is in order to avoid prioritizing areas solely representing ecotones, which may not represent species within their core ranges (**Figure 1.7**). That is, to avoid the problem of over-selecting minimum sets (Gaston et al. 2001). For conservation planning, nestedness implies conservation of the richest sites whereas turnover implies the conservation of several different sites (to incorporate as many different species as possible), though these may not be as species rich (Baselga 2010).

a)

g	g	g	e	g
w	w	w	w	w
w	w	e	w	w
w	w	w	w	w

b)

g	g	g	g	g
w	w	e	w	w
w	e	e	e	w
w	w	w	w	w

Figure 1.7 The shaded area represented in case b) might be a better choice for conserving the core range of species 'e' although the focal cell/area (unshaded) does not reflect high beta diversity due to species turnover, as in case a). High beta diversity due to species turnover does not necessarily translate to better coverage of a species' range.

In addition, disentangling nestedness and turnover components of beta diversity can also help in investigating the drivers of such beta diversity patterns. For example, spatial turnover may suggest barriers to dispersal or niche differences between faunas (Baselga 2010). At a large scale, this approach can help to identify regional endemic centres or, at a finer scale, places where species ranges are reflecting fragmentation.

All of the similarity/dissimilarity measures discussed so far to investigate beta diversity patterns refer to pair-wise comparisons between the focal sample and the neighbouring sample. There is some debate as to the use of multiple-site similarity measures over pair-wise comparison measures depending on the goal of the analysis (Diserud & Ødegaard 2007; Koleff et al. 2003). Some contend that if the goal of the study is to identify, *a priori*, how species composition changes along a specified gradient then pair-wise comparison is appropriate, but if the analysis involves looking at a selection of samples from a larger region (landscape), then multiple-site measures should be used (Diserud and Ødegaard 2007). It is argued that multiple-site similarity measures avoid lack of independence between pair-wise similarities due to repetition of each sample in several pair comparisons. Variants of such measures, that are additionally independent of richness, have been developed to allow comparison of more than two sites across a landscape (Baselga et al. 2007) e.g. M_{sim} (**Table 1.1**).

1.4 Choosing an appropriate Beta diversity measure

In using the distance approach to analyse beta diversity it is necessary to consider the similarity/dissimilarity measure used for the analysis since different measures may emphasize different aspects of species compositional change (see section 'Measuring Beta Diversity'). Past studies on beta diversity have been criticized for using only 'broad-sense' dissimilarity measures based mainly on the proportion of species shared between samples, and thus ignoring the contribution of species gained and species lost between samples, the essence of beta diversity (Koleff et al. 2003).

An important consideration in choosing a beta diversity measure is whether the index is applicable with incidence data (occurrence or presence/absence) or abundance data. Incidence based indices (because they essentially ignore the relative abundance of species in the sample) are sensitive to sample size especially in communities containing many rare species (Chao et al. 2005). If there are many rare species within the entire community, the rare species detected in one sample (say the focal cell) are more likely to be different than the rare species detected in the neighbour cell (Chao et al. 2005). Thus the similarity between the two sample areas is underestimated, which can lead to overestimation of beta diversity. It is likely that using the appropriate abundance-based beta diversity measure using abundance data (when it is available) is a better approach for beta diversity analysis. However, the availability of abundance data at large scales is usually not a feasible reality. Therefore, at least at present, most beta diversity analyses at large scales may be restricted to using presence/absence data.

Initially, this study will focus on the Simpson dissimilarity index (see **Table 1.1**) since it is less sensitive to differences in species richness (as discussed above), and is among the most widely used measures of turnover based on occurrence data (Olden et al. 2006).

1.5 Explaining Beta Diversity Patterns

Explaining the underlying drivers of any beta diversity patterns requires explaining how species assemblages arise. An area's species composition is a result of the behaviour of individuals of different species (Tuomisto and Ruokolainen 2006). This principle has led to the formulation of different hypotheses regarding the formation of community composition. One of these hypotheses is Hubbell's neutral theory which essentially states that species composition varies as a result of the random mortality and random dispersal that individuals of a species may be subject to (Tuomisto & Ruokolainen 2006). Neutral theory is based on the concept of functional equivalence which states that the *individuals* of species occupying similar trophic levels are equivalent in some measures including birth/death rate, dispersal and speciation and hence any species is equally likely to occupy a given space (Hubbell 2005). This hypothesis implies that species composition fluctuates randomly, but in a spatially auto-correlated way i.e. two sites which are closer to each other will tend to have more similar composition than two sites which are further apart. However, any spatial structure in species composition is due only to this autocorrelation and not as a result of variation in environmental variables (Tuomisto & Ruokolainen 2006).

The 'niche hypothesis' states that environmental conditions are important drivers of species composition as a result of habitat specialization by species occupying different habitats within a gradient (Whittaker 1972). If this is the case, then diversity patterns across different landscapes will vary differently and may be explained by specific environmental conditions or a suite of them (Tuomisto & Ruokolainen 2006).

A related scenario, arising from niche hypothesis, implies that species composition is driven by the ratio of dominant species to rare species (Tuomisto & Ruokolainen 2006). In this case, a suite of common species that are superior competitors come to dominate the landscape, resulting in community homogeneity (Pitman et al. 2001). The particular suite of dominant species (oligarchies) may vary across large landscapes but even at large spatial scales, composition will remain very similar as a result of the relatively greater proportion of dominant species to rare species (Tuomisto and Ruokolainen 2006). All these three explanations of community assemblage processes give rise to testable predictions regarding diversity patterns, which provide a basis for analysing beta diversity within the present study area.

From these scenarios, it is possible to make some predictions regarding the behaviour of beta diversity patterns. If the same dominant species always make up the greatest proportion of species in an assemblage (Pitman scenario), then species composition across the landscape remains regular, and hence beta diversity is low and similar over pairs of sites (Tuomisto & Ruokolainen 2006). Variation in beta diversity will be random and not explainable by geographic distance or environmental factors (Tuomisto & Ruokolainen 2006). Neutral theory states that species composition is the result of a random, spatially autocorrelated process (Tuomisto & Ruokolainen 2006). Therefore, neutral theory predicts that species turnover will exhibit spatial pattern beyond what can be explained by climate, but it will be uniform in all regions because it is not correlated to local conditions (Fattorini and Baselga 2011). Spatially closer sites will be more similar to each other in species composition than more distant sites (autocorrelation), but any resulting patterns in beta diversity are purely due to this spatial autocorrelation (Tuomisto &

Ruokolainen 2006). Lastly, the niche hypothesis predicts that species composition varies with variation in environmental conditions. Therefore turnover reflects ecological sorting and sites that are similar in environmental conditions will be more similar to each other (Fattorini & Baselga 2011). Hence beta diversity patterns will be explainable by environmental factors such as climatic conditions. All these predictions can be tested using a distance approach.

1.6 Analyzing Beta Diversity: Raw Data or Distance Approach?

Beta diversity itself is not an ecological entity but can be thought of as a distance measure based on the difference in species composition between sites. For the purposes of this study, ***beta diversity is defined as the amount of variation in species composition across sites within a landscape*** (Moreno 2001; Tuomisto and Ruokolainen 2006). This is to differentiate from alternative definitions that have variably defined beta diversity as the *rate* of change in species composition along spatial or environmental gradients (Vellend 2001), or change in species composition over *time* as a result of the displacement of some species by others (Olden et al. 2006).

Recent reviews have elucidated conceptual and statistical approaches to analysing beta diversity patterns depending on the ecological question being addressed (Anderson et al. 2011; Tuomisto and Ruokolainen 2006). Tuomisto & Ruokolainen (2006) make a distinction between explaining and analysing beta diversity. Using a hierarchical approach to define levels of abstraction, the authors explain that defining the variation in species composition (beta diversity) within a community is a level-2 question which requires level-1 or raw data (count/abundance data of species). This is called the 'raw data' approach. On the other hand, examining the variation in beta diversity, i.e., the variation in the variation in species composition is a level-3 question requiring level-2 data such as a distance matrix based on dissimilarity between all pairs of sites. This is termed the 'distance' approach.

Using Tuomisto & Ruokolainen (2006) definitions, research question 2 of this study is a level-3 question seeking to explain the variation in beta diversity, that is, the variation in the variation in species composition. This necessitates a distance approach and therefore application of standard distance methods to test statistical dependence of the response variable (in this case variation in beta diversity represented by a distance matrix of dissimilarity) to explanatory variables (geographic, environmental or climatic factors), such as multiple regression or generalized distance/dissimilarity modelling (GDM) (Anderson et al. 2011; Tuomisto and Ruokolainen 2006).

Linear matrix regression refers to analysis involving the correlation of two distance matrices where one matrix represents the dissimilarity in species composition between all pairs of a set of sites, and the other matrix represents distance between sites based on some environmental variable measured at the set of sites (e.g.) climatic variables (Ferrier et al. 2007). Regression can also be carried out on geographic distance between sites (spatial autocorrelation) using a measure of spatial distance (e.g.) Euclidean distance.

GDM is a variant of matrix regression designed to correct for two types of non-linearity common to

ecological data (Ferrier et al. 2007). One is the curvilinear relationship between environmental/spatial distance and compositional dissimilarity. This relationship reaches an eventual asymptote value of greatest dissimilarity, above which it does not increase regardless of increasing environmental distance, mostly due to the fact that most dissimilarity measure are bounded by 0 and 1 (Ferrier et al. 2007). The other type of non-linearity GDM is meant to correct refers to the rate of compositional turnover along an environmental gradient. Linear matrix regression assumes this rate to remain constant across any variables' range. GDM allows for changing rates of compositional turnover which is more reflective of real world data (Ferrier et al. 2007). Both types of non-linearity are likely to be present in large study areas encompassing long and varied environmental gradients (Ferrier et al. 2007), such as our study area.

An understanding of the processes driving observed beta diversity patterns also requires the application of null models, which begin with the simplest scenario (Hubbell 2005), in order to parse out patterns truly due to (e.g.) environmental variables (Anderson et al. 2011). Analysis of spatial autocorrelation or distance decay of similarity, which is the "hallmark" of neutral theory (Palmer 2005; Tuomisto and Ruokolainen 2006) can be used to test whether observed patterns in the measure of interest (beta-diversity) are just artefacts of spatial structuring or indeed indicate a centre of high beta diversity (Crisp et al. 2001).

Spatial autocorrelation can be defined as the property of two areas being more or less similar than expected by chance (with regard to some value e.g. a beta diversity measure) depending on the distance between them (Legendre 1993). In this case it would be expected that sites closer to each other would have more similar beta diversity values and sites further apart would have less similar beta diversity values. Autocorrelation is a property exhibited by all variables observed across time or space, including ecological ones. Spatial autocorrelation implies that the environment is spatially structured reflecting patterns of gradients or patchiness (Legendre 1993). This structuring is a functional factor giving rise to similar patterns in biological phenomena. Therefore, studying and accounting for spatial structuring within a system is an important factor in the study of any biogeographic patterns (Legendre 1993).

1.7 Thesis Focus

1.7.1 Knowledge Gaps

The preceding overview of the literature on beta diversity patterns reveals the debate about the processes that drive beta diversity patterns and the role of historic and environmental factors in explaining such patterns, and at what scale they operate. A focus on global and very local scale studies reveals a gap in the application of intermediate-scale studies to examine beta diversity patterns at the regional level. In addition, studies of beta diversity have been confined mostly to vertebrate species such as birds, mammals and amphibians. Understanding spatial patterns of tree diversity is especially important within the context of protecting tropical forests for carbon storage purposes (e.g. REDD+ initiatives), where these initiatives need to be used to maximize diversity conservation and not just focus on biomass. With this thesis, I attempt to bridge these knowledge gaps by examining beta diversity patterns of trees at a regional scale.

1.7.2 AIM & OBJECTIVES

This thesis aims to model and analyze spatial turnover patterns in tree species composition throughout Mesoamerica in order to inform conservation planning efforts within this biodiversity hotspot. I aim to achieve this by utilizing an increasingly recognized modelling approach known as Generalized Dissimilarity Modelling (GDM) to develop a spatial model of tree species turnover for the study region. I attempt to elucidate the importance of spatial and environmental factors in explaining species turnover patterns at both the modelling scale and the ecoregion unit scale. I apply my model, in combination with other biodiversity assessment approaches and the species-area relationship, to answer two specific conservation questions.

The specific objectives of this thesis are as follows:

Objective 1. Implement the GDM approach to predict patterns of compositional turnover in plant species across the Meoamerican region, in order to answer the following research questions:

- What variables best predict the variation in tree species composition patterns?
- How does the choice of a beta diversity index affect the model outcomes?
- How do tree species compositional patterns vary across the region compared with other regional classifications (e.g. WWF ecoreigons) or turnover patterns of other taxa?

Objective 2. Apply the GLOBIO framework and GDM approach to assess the impact of land use change on tree species diversity within the region, to answer the following questions:

- What proportion of habitat and species has been lost as a result of land use changes up to the year 2000?
- What is the effect of land use change on both areas that are compositionally more similar to other areas in the region (low beta diversity), and those that are less similar (high beta diversity)?
- How does habitat and species loss vary among ecoregions?

Objective 3. Assess the representativeness of the regional protected areas system in relation to species composition turnover patterns, by addressing the following questions:

- How representative of tree diversity is the PA system as a whole?
- Which ecoregions are best or least represented in the PA system?
- What environmental factors are associated with level of protection within the protected area system?

Objective 4. Evaluate whether the ecoregions classification system captures similarity in species composition (species turnover or beta diversity) among ecoregions.

- How does species compositional similarity vary among ecoregions?
- What is the contribution of spatial distance, mean annual temperature and mean annual precipitation as potential drivers of variation in species turnover among ecoregions?

1.8 Study Area

Mainland Mesoamerican encompasses the five southern states of Mexico (Campeche, Chiapas, Quintana Roo, Tabasco, and Yucatan) and all Central American countries (Belize, Guatemala, El Salvador, Honduras, Nicaragua, Costa Rica and Panama) (**Figure 1.8**). The region encompasses a land area of 768,990 km^2 and is home to a population of about 45 million (half of which is thought to live below the poverty line), which continues to grow at approximately 2% per year (Miller et al. 2001). Developmental pressures arising from poverty and an increasing population are drivers of the environmental degradation occurring over recent years (Conservation International 2004). The region is threatened by intense land use change (Golicher et al. 2012) including a deforestation rate of about 2.1% per year since measured in 1999 (Miller et al. 2001).

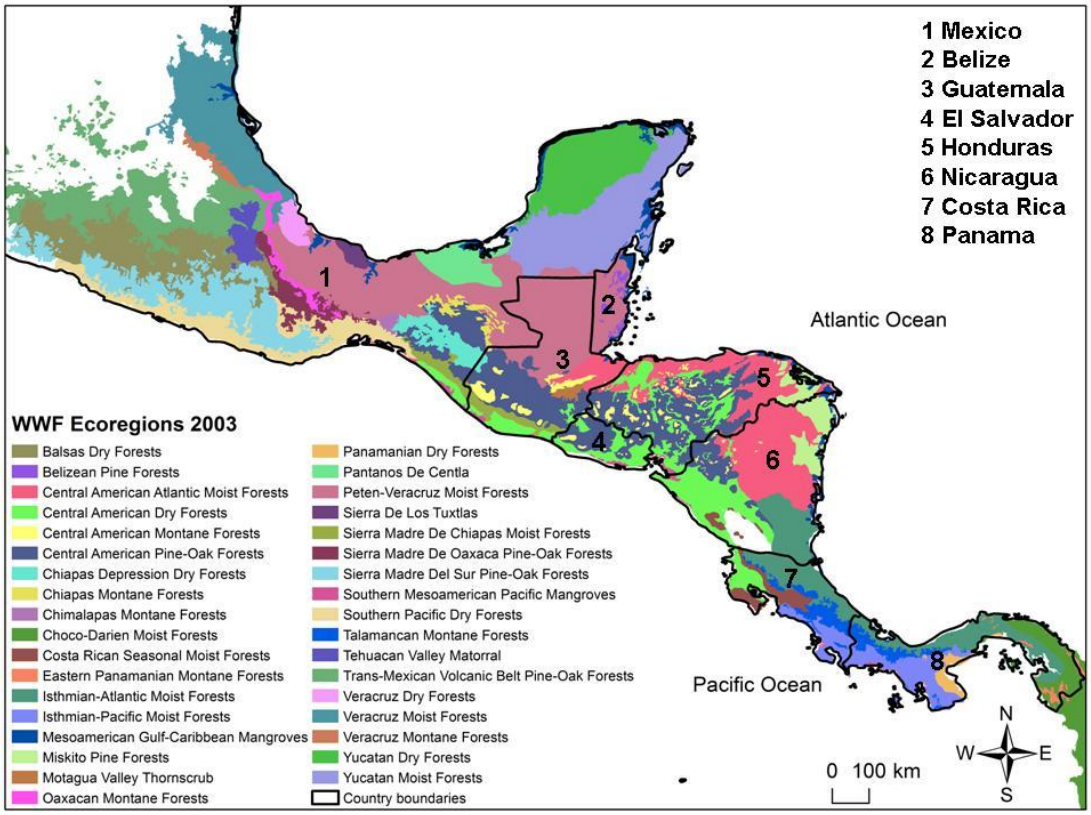


Figure 1.8 Map of Mesoamerican study region.

Mesoamerica is recognized as a centre of plant diversity and endemism, and as a world diversity hotspot (Sarkar et al. 2009). It holds between 7 to 10% of the planet's total known biodiversity (Miller et al. 2001) (Conservation International 2004). It is also estimated to hold 24,000 vascular plant species, 21% of which are endemic (Conservation International 2001). The region encompasses the Selva Maya forest

which is the largest, continuous area of tropical rainforest, second only to the Amazon (Conservation International 2004). Annual average temperature in the region varies between 7.5 ºC to 32.5ºC, and rainfall varies between 500 to 5000 mm per year (Miller et al. 2001; Sarkar et al. 2009). The region is expected to undergo significant changes in temperature, rainfall and seasonality under climate change scenarios and this is expected to seriously impact the region's biodiversity, especially in sensitive areas such as mountain and coastal habitats (Jankowski et al. 2009).

In recent years, the countries in the Mesoamerican region have made significant efforts to prioritize environmental issues, including conservation (Conservation International 2004). Since 1997, the Central American System of Protected Areas (SICAP) and Central American Commission on the Environment and Development (CCAD) have received political backing from all states within the region, signalling a regional approach to solving environmental issues (Conservation International 2001). One of the main goals of such regional cooperation is to protect forests and their biodiversity from threats posed by development challenges such as urban expansion, unsustainable agriculture, and uncontrolled logging (Conservation International 2004). In each country, a system of protected areas has been established. Throughout the region, this amounts to a total of 600 protected areas, representing about 20% of the terrestrial land mass. This is complemented by a system of biological corridors known as the Mesoamerican Biological Corridor (Conservation International 2004). However, it is important to highlight that two-thirds of the protected areas cover less than 10,000 hectares each; only about 38 areas protect more than 50,000 hectares (Conservation International 2001).

The region has been classified into 35 ecoregions including lowland and mountain forests, grasslands, pine savannahs, and mangroves (**Figure 1.8**). These were delineated based on known mapping of floristic and faunal distributions, vegetation types, and data and consultation from experts (Olson et al. 2001). These ecoregion units have been used as a basis for conservation planning efforts within Mesoamerica, such as in gap analyses to assess how much they are represented within existing protected area networks (Londoño-Murcia et al. 2010; Xiao-jun et al. 1999).

1.9 Thesis Structure

Chapter 1: Introduction and literature review

This chapter provides a comprehensive background on the concept of beta diversity, how it is measured, and what processes may explain beta diversity patterns. It also places the thesis and its focus on beta diversity within the context of ecological theory regarding community assemblage processes.

Chapter 2: Developing a spatial model of tree species compositional turnover using Generalized Dissimilarity Modelling

This chapter describes the development of the spatial model of tree beta diversity. I attempt to identify the most important predictors of tree beta diversity within study limitations. I discuss methodological limitations related to model development and evaluate model performance. I also discuss the resulting patterns of tree beta diversity within the context of other studies focusing on beta diversity patterns of other taxa.

Chapter 3: Incorporating beta diversity into estimates of tree diversity loss in the Mesoamerican region

This constitutes the first chapter applying the tree species spatial turnover model to specific conservation questions. In this chapter, I estimate the amount of habitat and species presumed lost up to a baseline year. I discuss the findings in relation to previous studies, and highlight conservation implications for the region.

Chapter 4: Assessing Representativeness of the Protected Area System

This chapter is the second application of the model. Here I assess how well the regional protected areas system captures tree diversity patterns, as represented by beta diversity. I highlight the variation in protection at this regional scale. I discuss the findings in relation to other assessments of protected areas representativeness.

Chapter 5: Assessing similarity of ecoregion units in Mesoamerica

This chapter examines beta diversity among ecoregions. The prevailing consensus is that ecoregions are equivalent units of diversity. The analysis showed overlap in species composition across ecoregions revealing the inherent hierarchical structure of species composition among them, and suggesting common historical processes for tree lineages in this region.

Chapter 6 Discussion

This section summarizes the findings of each chapter in relation to the original objectives of the thesis. I integrate and evaluate the major ecological and conservation implications of the study. I also discuss cross-cutting limitations throughout the thesis and list ideas for future work.

1.10 REFERENCES

Allnutt, T. F., Ferrier, S., Manion, G., Powell, G. V. N., Ricketts, T. H., Fisher, B. L., Harper, G. J., Irwin, M. E., Kremen, C., Labat, J.-N., Lees, D. C., Pearce, T. a. and Rakotondrainibe, F., 2008. A method for quantifying biodiversity loss and its application to a 50-year record of deforestation across Madagascar. *Conservation Letters*, *1* (4), 173–181.

Anderson, M. J., Crist, T. O., Chase, J. M., Vellend, M., Inouye, B. D., Freestone, A. L., Sanders, N. J., Cornell, H. V., Comita, L. S., Davies, K. F., Harrison, S. P., Kraft, N. J. B., Stegen, J. C. and Swenson, N. G., 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology Letters*, *14* , 19–28.

Baker, T. R., Jones, J. P. G., Rendón Thompson, O. R., Cuesta, R. M. R., Del Castillo, D., Aguilar, I. C., Torres, J. and Healey, J. R., 2010. How can ecologists help realise the potential of payments for carbon in tropical forest countries? *Journal of Applied Ecology*, *47* , 1159–1165.

Baselga, A., Jiménez-Valverde, A. and Niccolini, G., 2007. A multiple-site similarity measure independent of richness. *Biology Letters*, *3* (6), 642–5.

Baselga, A., 2010. Multiplicative partition of true diversity yields independent alpha and beta components; additive partition does not. *Ecology*, *91* (7), 1974–81.

Baselga, A., 2010. Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, *19* (1), 134–143.

Chao, A., Chazdon, R. L., Colwell, R. K. and Shen, T.-J., 2005. A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecology Letters*, *8* (2), 148–159.

Conservation International, 2001. *CEPF Ecosystem Profile: The Southern Region Of The Mesoamerica Biodiversity Hotspot*. Washington, D.C.

Conservation International, 2004. *CEPF Ecosystem Profile: Northern Region Of The Mesomaerica Biodiversity Hotspot*. *Analysis*. Washington, D.C.

Cowling, R. M. and Pressey, R. L., 2001. Rapid plant diversification: planning for an evolutionary future. *Proceedings of the National Academy of Sciences of the United States of America*, *98* (10), 5452–7.

Crisp, M. D., Laffan, S., Linder, H. P. and Monro, A., 2001. Endemism in the Australian flora. *Journal of Biogeography*, *28* (2), 183–198.

Crisp, M. D., Laffan, S., Linder, H. P. and Monro, 2002. Endemism in the Australian flora. *Journal of Biogeography*, *28* (2), 183–198.

Díaz, S., Hector, A. and Wardle, D. a, 2009. Biodiversity in forest carbon sequestration initiatives: not just a side benefit. *Current Opinion in Environmental Sustainability*, *1* (1), 55–60.

Dickson, B., Dunning, E., Killen, S., Miles, L. and Pettorelli, N., 2009. *Carbon markets and forest conservation: A review of the environmental benefits of REDD mechanisms.*

Diserud, O. H. and Ødegaard, F., 2007. A multiple-site similarity measure. *Biology Letters*, *3* (October 2006), 20–22.

Dunlop, M., Hilbert, D., Ferrier, S., House, A., Liedloff, A., Prober, S. M., Smyth, A., Martin, T. G., Harwood, T., Williams, K. J., Fletcher, C. and Murphy, H., 2012. *The implications of climate change for biodiversity conservation and the National Reserve System: Final Synthesis. A report prepared for the Department of Sustainability, Environment, Water, Population and Communities, and the Department of Climate Change*. Canberra.

Ebeling, J. and Yasué, M., 2008. Generating carbon finance through avoided deforestation and its potential to create climatic, conservation and human development benefits. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363* (1498), 1917–24.

Fairbanks, D., Reyers, B. and van Jaarsveld, A. S., 2001. Species and environment representation: selecting reserves for the retention of avian diversity in KwaZulu-Natal, South Africa. *Biological Conservation*, *98* (3), 365–379.

Faith, D., Ferrier, S. and Walker, P., 2004. The ED strategy: how species‐level surrogates indicate general biodiversity patterns through an "environmental diversity" perspective. *Journal of Biogeography*, *31* , 1207–1217.

Faith, D. P., Carter, G., Cassis, G., Ferrier, S. and Wilkie, L., 2003. Complementarity, biodiversity viability analysis, and policy-based algorithms for conservation. *Environmental Science & Policy*, *6* (3), 311–328.

Fattorini, S. and Baselga, A., 2011. Species richness and turnover patterns in European tenebrionid beetles. *Insect Conservation and Diversity*, 1–15.

Ferrier, S., Drielsma, M., Manion, G. and Watson, G., 2002. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales . II . Community-level modelling. *Biodiversity and Conservation*, *11* , 2309–2338.

Ferrier, S. and Drielsma, M., 2010. Synthesis of pattern and process in biodiversity conservation assessment: a flexible whole-landscape modelling framework. *Diversity and Distributions*, *16* (3), 386–402.

Ferrier, S., Manion, G., Elith, J. and Richardson, K., 2007.a Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions*, *13* (3), 252–264.

Ferrier, S., Manion, G., Elith, J. and Richardson, K., 2007.b Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions*, *13* (3), 252–264.

Ferrier, S., Powell, G. V. N., Richardson, K. S., Manion, G., M, J., Allnutt, T. F., Cameron, S. E., Mantle, K., Burgess, N. D., Faith, D. P., Lamoreux, J. F., Kier, G., Hijmans, R. J., Funk, V. A., Cassis, G. A., Fisher, L., Flemons, P., Lees, D., Lovett, J. C., Rompaey, R. S. A. R. Van, Overton, J. M., Allnutt, F., Fisher, B. L., Lovett, J. O. N. C. and Rompaey, R. S. A. R. V. A. N., 2004. Mapping More of Terrestrial Global Biodiversity for Conservation Assessment. *BioScience*, *54* (12), 1101–1109.

Ferrier, S., 2002. Mapping Spatial Pattern in Biodiversity for Regional Conservation Planning : Where to from Here ? *Systematic Biology*, *51* (2), 331–363.

Fitzpatrick, M. C., Sanders, N. J., Normand, S., Svenning, J.-C., Ferrier, S., Gove, A. D. and Dunn, R. R., 2013. Environmental and historical imprints on beta diversity: insights from variation in rates of species turnover along gradients. *Proceeding of The Royal Society of Biological Sciences*, *280* (1768), 1–8.

Galindo, L., de Miguel, C. and Ferrer, J., 2010. *Vital Climate Change Graphics for Latin America and the Caribbean*.

Gasner, M. R., Jankowski, J. E., Ciecka, A. L., Kyle, K. O. and Rabenold, K. N., 2010. Projecting the local impacts of climate change on a Central American montane avian community. *Biological Conservation*, *143* (5), 1250–1258.

Gaston, K. J., Rodrigues, a. S. L., van Rensburg, B. J., Koleff, P. and Chown, S. L., 2001. Complementary representation and zones of ecological transition. *Ecology Letters*, *4* (1), 4–9.

Gering, J. C. and Crist, T. O., 2002. The alpha-beta-regional relationship: providing new insights into local-regional patterns of species richness and scale dependence of diversity components. *Ecology Letters*, *5* (3), 433–444.

Gitay, H., Suarez, A. V, Watson, T. and Dokken, D. J., 2002. *Climate change and biodiversity. IPCC Technical Paper V. Climate Change and Biodiversity*

Golicher, D. J., Cayuela, L. and Newton, A. C., 2012. Effects of Climate Change on the Potential Species Richness of Mesoamerican Forests. *Biotropica*, *44* (3), 284–293.

González-Espinosa, M., María Rey-Benayas, J., Ramírez-Marcial, N., Huston, M. a. and Golicher, D., 2004. Tree diversity in the northern Neotropics: regional patterns in highly diverse Chiapas, Mexico. *Ecography*, *27* (6), 741–756.

Grainger, A., Boucher, D. H., Frumhoff, P. C., Laurance, W. F., Lovejoy, T., McNeely, J., Niekisch, M., Raven, P., Sodhi, N. S., Venter, O. and Pimm, S. L., 2009. Biodiversity and REDD at Copenhagen. *Current Biology*, *19* (21), r974–976.

Harvey, C. a., Dickson, B. and Kormos, C., 2010. Opportunities for achieving biodiversity conservation through REDD. *Conservation Letters*, *3* (1), 53–61.

Hole, D. G., Willis, S. G., Pain, D. J., Fishpool, L. D., Butchart, S. H. M., Collingham, Y. C., Rahbek, C. and Huntley, B., 2009. Projected impacts of climate change on a continent-wide protected area network. *Ecology Letters*, *12* (5), 420–31.

Hubbell, S. P., 2005. Neutral theory in community ecology and the hypothesis of functional equivalence. *Functional Ecology*, (1917), 166–172.

Jankowski, J. E., Ciecka, A. L., Meyer, N. Y. and Rabenold, K. N., 2009. Beta diversity along environmental gradients: implications of habitat specialization in tropical montane landscapes. *The Journal of Animal Ecology*, *78* (2), 315–27.

Jost, L., DeVries, P., Walla, T., Greeney, H., Chao, A. and Ricotta, C., 2010. Partitioning diversity for conservation analyses. *Diversity and Distributions*, *16* (1), 65–76.

Jurasinski, G., Retzer, V. and Beierkuhnlein, C., 2009. Inventory, differentiation, and proportional diversity: a consistent terminology for quantifying species diversity. *Oecologia*, *159* (1), 15–26.

Koleff, P., Gaston, K. J. and Lennon, J. J., 2003. Measuring beta diversity for presence-absence data. *Journal of Animal Ecology*, *72* (3), 367–382.

Koleff, P. and Gaston, K. J., 2002. The relationships between local and regional species richness and spatial turnover. *Global Ecology and Biogeography*, *11* (5), 363–375.

Kraft, N. J. B., Comita, L. S., Chase, J. M., Sanders, N. J., Swenson, N. G., Crist, T. O., Stegen, J. C., Vellend, M., Boyle, B., Anderson, M. J., Cornell, H. V., Davies, K. F., Freestone, a. L., Inouye, B. D., Harrison, S. P. and Myers, J. a., 2011. Disentangling the Drivers of Diversity Along Latitudinal and Elevational Gradients. *Science*, *333* (6050), 1755–1758.

Laffan, S. W. and Crisp, M. D., 2003. Assessing endemism at multiple spatial scales, with an example from the Australian vascular flora. *Journal of Biogeography*, *30* (4), 511–520.

Leathwick, J. R., Moilanen, a., Ferrier, S. and Julian, K., 2010. Complementarity-based conservation prioritization using a community classification, and its application to riverine ecosystems. *Biological Conservation*, *143* (4), 984–991.

Legendre, P., 1993. Spatial Autocorrelation : Trouble or New Paradigm ? *Ecology*, *74* (6), 1659–1673.

Lera Miles and Valerie Kapos, 2008. Reducing Greenhouse Gas Emissions from Deforestation and Forest Degradation: Global Land-Use Implications. *Science*, *320* , 1454–1455.

Londoño-Murcia, M. C., Tellez-Valdés, O. and Sánchez-Cordero, V., 2010. Environmental heterogeneity of World Wildlife Fund for Nature ecoregions and implications for conservation in Neotropical biodiversity hotspots. *Environmental Conservation*, *37* (02), 116–127.

Malhi, Y. and Grace, J., 2000. Tropical forests and atmospheric carbon dioxide. *Trends in Ecology & Evolution*, *15* (8), 332–337.

Margules, C. R. and Pressey, R. L., 2000. Systematic conservation planning. *Nature*, *405* (6783), 243–53.

McKnight, M. W., White, P. S., McDonald, R. I., Lamoreux, J. F., Sechrest, W., Ridgely, R. S. and Stuart, S. N., 2007. Putting beta-diversity on the map: broad-scale congruence and coincidence in the extremes. *PLoS Biology*, *5* (10), 2424–2432.

Melo, A. S., Rangel, T. F. L. V. B. and Diniz-Filho, J. A. F., 2009. Environmental drivers of beta-diversity patterns in New-World birds and mammals. *Ecography*, *32* (2), 226–236.

Midgley, G. F., Bond, W. J., Kapos, V., Ravilious, C., Scharlemann, J. P. and Woodward, F. I., 2010. Terrestrial carbon stocks and biodiversity: key knowledge gaps and some policy implications. *Current Opinion in Environmental Sustainability*, 264–270.

Miller, K., Chang, E. and Johnson, N., 2001. *Defining common ground for the Mesoamerican biological corridor. Management*. Washington, D.C.: World Resources Institute.

Mokany, K., Harwood, T. D., Overton, J. M., Barker, G. M. and Ferrier, S., 2011. Combining α - and β - diversity models to fill gaps in our knowledge of biodiversity. *Ecology Letters*, *14* , 1043–1051.

Moreno, C. E., 2001. *Métodos para Medir la Biodiversidad Director*. Zaragoza, España: M&T-Manuales y Tesis SEA, Vol 1.

Olden, J. D., Rooney, T. P. and N, P. S., 2006. On defining and quantifying biotic homogenization. *Global Ecology and Biogeography*, *15* , 113–120.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'amico, J. a., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P. and Kassem, K. R., 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience*, *51* (11), 933.

Palmer, M. W., 2005. Distance decay in an old-growth neotropical forest. *Journal of Vegetation Science*, *16* (2), 161–166.

Phelps, J., Guerrero, M. C., Dalabajan, D. a., Young, B. and Webb, E. L., 2010. What makes a "REDD" country? *Global Environmental Change*, *20* (2), 322–332.

Pitman, N. C. A., Terborgh, J. W., Silman, M. R., V, P. N., A, D., Cerón, C. E., Palacios, W. A. and Aulestia, M., 2001. Dominance and Distribution of Tree Species in Upper Amazonian Terra Firme Forests. *Ecology*, *82* (8), 2101–2117.

Pressey, R. L., Cabeza, M., Watts, M. E., Cowling, R. M. and Wilson, K. A., 2007. Conservation planning in a changing world. *Trends in Ecology & Evolution*, *22* (11), 583–92.

Rodríguez, J. P., Beard, T. D., Bennett, E. M., Cumming, G. S., Cork, S. J., Agard, J., Dobson, A. P. and Peterson, G. D., 2006. Trade-offs across Space , Time , and Ecosystem Services. *Ecology And Society*, *11* (1), 1–14.

Sarkar, S., Sánchez-Cordero, V., Londoño, M. C. and Fuller, T., 2009. Systematic conservation assessment for the Mesoamerica, Chocó, and Tropical Andes biodiversity hotspots: a preliminary analysis. *Biodiversity and Conservation*, *18* (7), 1793–1828.

Sesnie, S. E., Finegan, B., Gessler, P. E. and Ramos, Z., 2009. Landscape-Scale Environmental and Floristic Variation in Costa Rican Old-Growth Rain Forest Remnants. *Biotropica*, *41* (1), 16–26.

Steffan-Dewenter, I., Kessler, M., Barkmann, J., Bos, M. M., Buchori, D., Erasmi, S., Faust, H., Gerold, G., Glenk, K., Gradstein, S. R., Guhardja, E., Harteveld, M., Hertel, D., Höhn, P., Kappas, M., Köhler, S., Leuschner, C., Maertens, M., Marggraf, R., Migge-Kleian, S., Mogea, J., Pitopang, R., Schaefer, M., Schwarze, S., Sporn, S. G., Steingrebe, A., Tjitrosoedirdjo, S. S., Tjitrosoemito, S., Twele, A., Weber, R., Woltmann, L., Zeller, M. and Tscharntke, T., 2007. Tradeoffs between income, biodiversity, and ecosystem functioning during tropical rainforest conversion and agroforestry intensification. *Proceedings of the National Academy of Sciences of the United States of America*, *104* (12), 4973–4978.

Steinitz, O., Heller, J., Tsoar, A., Rotem, D. and Kadmon, R., 2005. Predicting Regional Patterns of Similarity in Species Composition for Conservation Planning. *Conservation Biology*, *19* (6), 1978–1988.

Steinitz, O., Heller, J., Tsoar, A., Rotem, D. and Kadmon, R., 2006. Environment, dispersal and patterns of species similarity. *Journal of Biogeography*, *33* (6), 1044–1054.

Strassburg, B. B. N., Kelly, A., Balmford, A., Davies, R. G., Gibbs, H. K., Lovett, A., Miles, L., Orme, C. D. L., Price, J., Turner, R. K. and Rodrigues, A. S. L., 2010. Global congruence of carbon storage and biodiversity in terrestrial ecosystems. *Conservation Letters*, *3* (2), 98–105.

Tuomisto, H. and Ruokolainen, K., 2006. Analyzing or explaining beta diversity? Understanding the targets of different methods of analysis. *Ecology*, *87* (11), 2697–2708.

Tuomisto, H., 2010. A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography*, *33* (1), 2–22.

Veech, J. a. and Crist, T. O., 2007. Habitat and climate heterogeneity maintain beta-diversity of birds among landscapes within ecoregions. *Global Ecology and Biogeography*, *16* (5), 650–656.

Veech, J. a., Summerville, K. S., Crist, T. O. and Gering, J. C., 2002. The additive partitioning of species diversity: recent revival of an old idea. *Oikos*, *99* (1), 3–9.

Vellend, M., 2001. Do commonly used indices of β-diversity measure species turnover? *Journal of Vegetation Science*, *12* (4), 545–552.

Venter, O., Laurance, W. F., Iwamura, T., Wilson, K. A., Fuller, R. A. and Possingham, H. P., 2009. Harnessing carbon payments to protect biodiversity. *Science*, *326* , 1368.

Whittaker, R. H., 1960. Vegetation of the Siskiyou Mountains , Oregon and California. *Ecological Monographs*, *30* (3), 279–338.

Whittaker, R. H., 1972. Evolution and Measurement of Species Diversity. *Taxon*, *21* (2/3), 213–251.

Wiersma, Y. F. and Urban, D. L., 2005. Beta Diversity and Nature Reserve System Design in the Yukon, Canada. *Conservation Biology*, *19* (4), 1262–1272.

Williams, K. J., Ferrier, S., Rosauer, D., Yeates, D., Manion, G., Harwood, T., Stein, J., Faith, D. P., Laity, T. and Whalen, A., 2010. *Harnessing Continent-Wide Biodiversity Datasets for Prioritising National Conservation Investment*. *Communities*. Canberra, Australia.

Wilson, K. a, Underwood, E. C., Morrison, S. a, Klausmeyer, K. R., Murdoch, W. W., Reyers, B., Wardell-Johnson, G., Marquet, P. a, Rundel, P. W., McBride, M. F., Pressey, R. L., Bode, M., Hoekstra, J. M., Andelman, S., Looker, M., Rondinini, C., Kareiva, P., Shaw, M. R. and Possingham, H. P., 2007. Conserving biodiversity efficiently: what to do, where, and when. *PLoS Biology*, *5* (9), 1850–1861.

Wu, F., Yang, X. J. and Yang, J. X., 2010. Additive diversity partitioning as a guide to regional montane reserve design in Asia: an example from Yunnan Province, China. *Diversity and Distributions*, *16* (6), 1022–1033.

Xiao-jun, L., Boucher, T. and Sayre, R., 1999. A conservation assessment of Central American vegetation and ecoregions. *Journal of Natural Resources*, *4* (4), 300–306.

2 CHAPTER 2

Developing a spatial model of tree species compositional turnover using Generalized Dissimilarity Modelling

2.1 INTRODUCTION

Tropical tree species diversity is an important indicator of overall conservation value as it correlates with other measures of biological importance, such as bird and insect species richness, and ecosystem productivity (Golicher et al. 2011). Understanding tree diversity is crucial to understanding tropical forest biodiversity since trees are providers of habitat, refuge and food for many other species (Cayuela, Gálvez-Bravo, et al. 2012). Therefore comparing different areas based on their tree diversity can be a good proxy for assessing their conservation priority. The establishment of representative conservation areas, then, depends on the degree of species compositional heterogeneity, or beta diversity, within the region of interest (Wu et al. 2010).

The need to incorporate environmental heterogeneity in conservation planning has been previously recognized (Cowling & Pressey 2001; Londoño-Murcia et al. 2010); but, the opportunity to incorporate beta diversity as a measure towards this is only recently being explored. The concept of beta diversity is receiving growing research attention, and there is a growing body of literature exploring the application of beta diversity in the context of conservation and climate change planning (Allnutt et al. 2008; Dunlop et al. 2012; Fairbanks et al. 2001; Faith et al. 2003, 2004; Ferrier and Drielsma 2010; Ferrier 2002; Ferrier et al. 2002, 2004, 2007; Fitzpatrick et al. 2013; Kraft et al. 2011; Leathwick et al. 2010; Mokany et al. 2011; Steinitz et al. 2005; Wiersma and Urban 2005; Williams et al. 2010; Wu et al. 2010).

For example, Jennings et al. (2008) proposes an approach for quantifying the "compositional representativeness" of fish communities between different freshwater ecoregions. Within this framework, the approach differentiates between representativeness and distinctiveness. Representativeness refers to the proportion of shared species in pair-wise comparisons between regions and distinctiveness refers to the proportion of unshared species. In this way, representativeness equates to the matching species component of beta diversity and distinctiveness equates to the species gain and loss components of beta diversity as discussed in Koleff et al. (2003). Other studies have developed approaches for linking beta diversity to reserve selection in order to maximize complementarity within a set of protected areas (Faith et al. 2003; Faith et al. 2004). These studies combine measures of beta diversity with ordination methods to estimate "environmental diversity (ED)" values for a given set of protected areas (Faith et al. 2003). Measures of beta diversity (such as Bray-Curtis similarity values) are used to create the ordination space over which ED values can be estimated for different sets of protected areas. The ED value essentially measures the amount of gaps in species representation for a given set of areas. The set of areas with the lowest ED value therefore best represents species diversity of the entire area (Faith et al. 2004). Arponen et al. (2008) developed a novel

method for selecting conservation sites based on modelling species richness and beta diversity. This "maximization of complementary richness" method finds the optimal set of sites by comparing the total fraction of species covered by different sets of sites (Arponen et al. 2008). The approach combines modelled species richness produced using generalized additive models with modelled beta diversity produced using generalized dissimilarity modelling.

Modelling spatial turnover in species composition or beta diversity represents an approach known as community-level modelling. This approach combines data from multiple species to produce information on spatial patterns of biodiversity at a collective community level rather than the individual species level (Ferrier & Guisan 2006). Modelling collective community properties of biodiversity such as species richness or beta diversity can be particularly useful for analyses utilizing large numbers of species, especially where a sizeable proportion of these species is rarely recorded (Ferrier & Guisan 2006), or for those utilizing presence-only data sets in which some species are poorly represented (Elith & Leathwick 2007), such as . In community-level modelling, information on data-rich species can help to inform models of data-poor species (Elith & Leathwick 2007), since this approach utilizes all available data across all species (Ferrier & Guisan 2006). In addition, such multi-response models that use the signal from several species in selecting predictor variables may be more robust for prediction because predictors are included because of their strong signal across all species (Elith & Leathwick 2007).

The most well known application of community-level modelling is in modelling species richness (Benito et al. 2013; Colwell and Coddington 1994; Golicher et al. 2012; Harte et al. 2009; Longino et al. 2002; Walther and Morand 1998). One way by which building predictive models of beta diversity is being explored is with the development of a method called Generalized Dissimilarity Modelling (GDM). GDM is a statistical technique developed to predict spatial patterns of turnover in community composition (beta diversity) across large scales (Ferrier et al. 2007). The method is based on the principle of the distance-decay of similarity between any two areas (Morlon et al. 2008; Nekola and White 1999; Soininen et al. 2007). As the geographic or ecological distance between any two areas increases, their similarity in species composition decreases or decays (Faith et al. 1987; Nekola and White 1999). In other words, as ecological distance increases, dissimilarity in species composition also increases (Faith et al. 1987). GDM models this relationship between ecological distance and dissimilarity in species composition.

In doing so, GDM accounts for two types of non-linearity frequently encountered in ecological data (Ferrier et al. 2007). The first of these refers to the measures of compositional dissimilarity utilized (see Koleff et al. 2003 for a good review). These measures are typically bounded by 0 and 1. As ecological distance between two areas increases until they share no species, the dissimilarity value reaches an asymptotic value of 1. This curvilinear relationship is especially relevant when considering a large study area that is environmentally heterogeneous (Ferrier et al. 2007), where a large proportion of compared sites may share no species, as is the case in our study area. The second case of non-linearity refers to the issue of non-stationarity or variation in the rate of compositional turnover along ecological gradients (Ferrier et al. 2007). GDM addresses this

issue by fitting non-linear spline functions to the raw ecological variables themselves, rather than to the pair-wise distances derived from these variables (Ferrier et al. 2007). These spline functions therefore model the relationship between a particular variable and compositional dissimilarity, as well as how this relation varies along the gradient of that particular variable. In this way, GDM allows for a separate non-linear relationship between compositional dissimilarity and ecological distance for each ecological variable (Ferrier et al. 2007; Rosauer et al. 2014). These two types of non-linearity are likely to be common in large study areas that encompass different habitats and therefore span long environmental gradients (Ferrier et al. 2007), as is the case in the Mesoamerican study area. A final GDM model summarizes the relationship between compositional dissimilarity and ecological distance across all variables to predict dissimilarity between sites, and represent the overall spatial pattern in species compositional turnover (Rosauer et al. 2014).

The flexibility of the GDM approach in modelling compositional turnover based on measured environmental variables had led to its increasing application in biodiversity assessments (Ferrier et al. 2007). For example, it has been used to classify areas of similar biological composition (Ferrier et al. 2007), identify biological collection survey gaps (Ferrier 2002; Williams et al. 2010), and assess the representativeness of protected area systems (Ferrier et al. 2004; Williams et al. 2010).

The objective of this chapter is to implement the GDM approach to predict patterns of compositional turnover in plant species across the Meoamerican region, in order to answer the following research questions:

- How does the choice of a beta diversity index affect the model outcomes?
- What variables best predict the variation in tree species composition patterns?
- How do tree species compositional patterns vary across the region compared with other regional classifications (e.g. WWF ecoreigons) or turnover patterns of other taxa?

2.2 METHODS

The GDM method was used to predict spatial patterns of turnover in tree species composition across the Mesoamerican region. A GDM was fitted to species dissimilarity data based on georeferenced specimen records from the Global Biodiversity Information Facility (GBIF) database. Predicted species turnover from the model was then used to identify areas having low and high similarity to the rest of the region. Results were also used to classify the region into areas according to species compositional similarity and compared to other biogeographic regionalization efforts, as detailed below.

2.2.1 Study area

Mainland Mesoamerican encompasses the five southern states of Mexico (Campeche, Chiapas, Quintana Roo, Tabasco, and Yucatan) and all Central American countries (Belize, Guatemala, El Salvador, Honduras, Nicaragua, Costa Rica and Panama). The region encompasses a land area of 768,990 km^2 and is home to a population of about 45 million (half of which is thought to live below the poverty line), which continues to grow at approximately 2% per year (Miller et al. 2001). Developmental pressures arising from poverty and an increasing population are drivers of the environmental degradation occurring over recent years (Conservation International 2004). The region is threatened by intense land use change (Golicher et al. 2012) including a deforestation rate of about 2.1% per year since measured in 1999 (Miller et al. 2001).

2.2.2 Species data

Tropical tree species diversity is an important indicator of overall conservation value as it correlates with other measures of biological importance, such as bird and insect species richness, and ecosystem productivity (Golicher et al. 2012). Understanding tree diversity is crucial to understanding tropical forest biodiversity since trees are providers of habitat, refuge and food for many other species (Cayuela, Gálvez-Bravo, et al. 2012). Therefore comparing different areas based on their tree diversity can be a good proxy for assessing their conservation priority.

Available species data is usually a composite of ad-hoc observations or of many data collection efforts that employ different methodologies and have been implemented at various spatial scales (Elith & Leathwick 2007). In the tropics, systematically collected species inventory data is sparse (Feeley and Silman 2011). This absence of data means that diversity studies on a regional scale must rely on data from museum and herbaria collections (Feeley and Silman 2011; Golicher et al. 2012). Recent data sharing efforts have aimed at collating such information and making it accessible globally. The most well-known of such efforts is the Global Biodiversity Information Facility (GBIF, www.gbif.org), which has collated millions of data entries for many known taxa from natural history collections, library materials, and databases (Yesson et al. 2007).

For this analysis, a species list for the region was collated based on species identified within inventory plots across the region as part of the Tree Biodiversity Network (BIOTREE-NET) initiative (Cayuela et al. 2012a).

There were a total of 2001 inventory plots across the region. Within these plots, mainly angiosperms were reported as woody trees and shrubs (height > 4m; most 2.5-10cm dbh). The ecoregions with the least number of plots, having less than 10 plots each were: Yucatan Dry Forests, Southern Mesoamerican Pacific Pacific Mangrove Forests, Belizean Pine Forests, Veracruz Moist Forests, Sierra Madre del Sure Pine-Oak Forests, Eastern Panamanian Montane Forests, Miskito Pine Forests, Sierra de lot Tuxtlas. The ecoregion with the highest number of plots was Central American Pine-Oak Forests followed by Isthmian Atlantic Moist Forests. No plots were registered in Guatemala since no information was obtained from that country.

A list was created for all angiosperm species recorded within all inventory plots. This species list was then used to download geo-referenced specimen data from GBIF for the entire study region. The dataset underwent taxonomic standardization using the R package TAXONSTAND Version 1.2, which standardizes plant names using The Plant List (http://www.theplantlist.org) to retrieve information about each species' taxonomic status (Cayuela et al. 2012b). The Plant List is a collaborative effort among botanic gardens and draws on information about plant nomenclature from various sources. For angiosperms, the Plant List assigns accepted genera to the families recognised by the Angiosperm Phylogeny Group according to the 2009 updated family and subclass level classifications (Chase and Reveal 2009). Standardization of the dataset showed that 50.7% of species names in the dataset were accepted names, 11.2% of names were synonyms, 1.7% had unresolved names, and 36.4% had names for which a determination could not be made on accepted, synonym or unresolved (Cayuela et al. 2012b).

In addition, the species data set was filtered to remove obvious duplicate records, records which had unknowns at any of the taxonomic levels (family, genus, species), records with coordinates rounded to the nearest degree, records with uncertainty values of > 10,000 meters where available, and records of any orchids, grasses or sedges. Finally, only species with more than ten presence records were included in the dataset. All processing was done in R v. 2.13.2 (R Development Core Team 2011).

It is important to note that an absence of inventory plots in an area (e.g. in Guatemala) does not imply that no species data was available for that area in this analysis. This is because the plots were only used to collate a species list for the region that was then used to download specimen record data from across the entire study region. The final dataset used for this analysis consisted of over 435,000 geo-referenced specimen records from GBIF, representing 2,561 species of mostly woody trees and shrubs.

2.2.3 Climate data

Both temperature and water are variables that directly affect plant growth, and have been used to explain patterns of species richness (Pausas and Austin 2001). Therefore these same variables may help to explain patterns of species turnover. Bioclimatic predictor grids and a digital elevation model (DEM) grid in raster format were derived from the WorldClim bioclimatic variables dataset (Hijmans et al. 2005). These predictors represent averages, extremes, and seasonal patterns in temperature and precipitation for the period 1950-2000 (Hijmans et al. 2005). Bioclimatic variable grids were downloaded from Worldclim

(www.worldclim.org) at 30 arc-seconds (~1 km) resolution and re-sampled to a resolution of 0.1 degree (~10 km). A new raster grid for each bioclimatic variable was derived by averaging the values of all pixels within the 30 arc-second bioclimatic grids to the desired 0.1 degree resolution. An actual evapotranspiration raster grid was also derived at 0.1 degree resolution from the MODIS global terrestrial evapotranspiration data set (MOD16) (Mu et al. 2011) covering the period 2000-2010. **Table 2.1** shows all climatic data used in the GDM model.

Table 2.1 List of all predictors used in GDM model. Bioclimatic and DEM predictor grids derived from Hijmans et al. (2005) and AET from Mu et al. (2011).

Climatic variable	Units / Calculation (if applicable)
Elevation (DEM =Digital Elevation Model)	Meters (m)
BIO1 = Annual Mean Temperature	^{o}C * 10
BIO2 = Mean Diurnal Range	(Mean of monthly (max temp - min temp))
BIO3 = Isothermality	(BIO2/BIO7) (* 100)
BIO4 = Temperature Seasonality	(standard deviation *100)
BIO5 = Max Temperature of Warmest Month	^{o}C * 10
BIO6 = Min Temperature of Coldest Month	^{o}C * 10
BIO7 = Temperature Annual Range	(BIO5-BIO6)
BIO8 = Mean Temperature of Wettest Quarter	^{o}C * 10
BIO9 = Mean Temperature of Driest Quarter	^{o}C * 10
BIO10 = Mean Temperature of Warmest Quarter	^{o}C * 10
BIO11 = Mean Temperature of Coldest Quarter	^{o}C * 10
BIO12 = Annual Precipitation	Millimetres (mm)
BIO13 = Precipitation of Wettest Month	Millimetres (mm)
BIO14 = Precipitation of Driest Month	Millimetres (mm)
BIO15 = Precipitation Seasonality	(Coefficient of Variation)
BIO16 = Precipitation of Wettest Quarter	Millimetres (mm)
BIO17 = Precipitation of Driest Quarter	Millimetres (mm)
BIO18 = Precipitation of Warmest Quarter	Millimetres (mm)
BIO19 = Precipitation of Coldest Quarter	Millimetres (mm)
AET = Actual Evapotranspiration	Mm/year
Geographic Distance	Decimal degrees; Lat/Long coordinate reference system; datum=WGS 1984

2.2.4 Creating a species GDM model

Generating a plant species GDM model involved four steps. First an appropriate beta diversity index was chosen. Secondly, a site-pairs dissimilarity matrix was generated as input to the GD Modeller software (Manion 2009). Thirdly, the GD Modeller software was used to generate spline functions for predictors to fit the model, according to the procedure described in Ferrier et al. (2007). Finally stepwise backward elimination was used to remove predictors that contribute minimally to the model. These steps are subsequently described in more detail.

2.2.4.1 Beta diversity index

The level of compositional dissimilarity or beta diversity between two sites can be measured using a variety of indices (Koleff et al. 2003). GDM allows for the use of any of these indices in its implementation. However, two properties of any index need to be considered when deciding which dissimilarity index to use, namely, sensitivity to under-sampling and sensitivity to richness differences between sites. Cardoso et al. (2009) evaluated the sensitivity of certain beta diversity indices to under-sampling or sampling effort. In that author's analysis of empirical datasets most of the beta diversity indices revealed true beta diversity as sampling effort increased, namely as a greater percentage of the species pool was sampled. In this case, however, it is not known how much of the actual species pool across the region the dataset used here has actually sampled, as it consists of occurrence records. Therefore, a measure that is robust to under-sampling is needed.

Cardoso et al. (2009) argued that certain indices were less likely to overestimate beta diversity when the common problem of under-sampling was present. Such indices used the smaller of the gain or loss components when calculating pair-wise beta diversity (Cardoso et al. 2009) (see **Chapter 1**, **Table 1.1** for indices and explanation). The B_{sim} index did not fulfill this criteria in the formulation used in Cardoso et al. (2009). However, the B_{sim} formulation provided by Koleff et al. (2003) did use the smaller of the mismatching components and this prevents the index from overestimating beta diversity if under-sampling were present. In this analysis, the formulation of the B_{sim} index provided by Koleff et al. (2003) was used to calculate pair-wise beta diversity between sample sites.

In addition, the B_{sim} index is insensitive to differences in species richness between sites being compared (Cardoso et al. 2009; Keil et al. 2012), unlike other measures such as B_{sor} that misrepresent differences in species richness as differences in beta diversity. This is important in the case of nestedness. Consider two sites each with a different number of species, where the total set of species in one site is a subset of the total set of species in the other site. Indices such as B_{sor} do not reflect this nestedness and assign a beta diversity value above zero, thereby reflecting the species richness difference. The B_{sim} index, however, makes the distinction between nestedness and true species turnover—gain or loss of species when comparing one site to another (Koleff et al. 2003).

For this reason the B_{sim} index was used in this analysis. Additionally and for comparison, the B_{sor} index was also implemented due to its wide use in beta diversity studies (Benito et al. 2013; Fitzpatrick et al. 2013; McCafferty 1998; Mokany et al. 2011; Swenson et al. 2010). It is important to note that B_{sor} is a measure of similarity whereas B_{sim} is a measure of dissimilarity. In this case, B_{sor} was converted to its dissimilarity equivalent (1- B_{sor}) in order for comparison to be made with B_{sim}. I compared the effect of using these different indices on the GDM outcomes.

2.2.4.2 Data sampling and generating a site-pairs dissimilarity matrix

The grid cell size of our study was defined at 0.1 degrees. The choice of resolution in studies such as this one is often a trade-off between sampling enough detail of all relevant environmental gradients for model purposes, and trying to compensate for sampling effort by excluding severely under-sampled sites. In this case, using under-sampled sites or grid cells, such as those reporting only one species, would produce erroneous predictions of high dissimilarity between sites. This would not be reflective of actual dissimilarity between sites but of differences in sampling effort across those sites. Therefore, it was necessary to make sure that only well-sampled grid cells were used for training the model. For this analysis, five different grid resolutions were explored to determine a resolution that would provide a balance between detail and sampling effort (**Table 2.2**). Based on this analysis, a decision was made to use a grid cell resolution of 0.1 degrees and to utilize only those grid cells (sites) with five or more species.

Table 2.2 List of grid resolutions trialled to determine study resolution and threshold for number of species per site.

Grid resolution (decimal degrees)	# occupied cells	# cells with >10 species	Percent coverage of occupied cells (%)	Mean number of species/cell
0.05	7931	4756	60	26
0.1	6777	3425	51	32
0.25	1820	1398	77	76
0.5	568	514	90	172
1.0	176	160	91	344

There were 9,544 grid cells covering the study area of which 6,777 cells or sites had species data (71% coverage). For *n* sites, this gave a total of $n(n-1)/2$ possible site-pair combinations (> 22 million). Computer memory constraints made it impractical to use all site-pair combinations to train the GDM model. The GD Modeller software (Manion 2009) handled a maximum of 1 million site-pair comparisons, depending on other parameters such as the number of predictors used. Also, inclusion of all possible site-pairs could bias results if sample sites were biased in geographic or environmental space in relation to the distribution of the study taxa (Rosauer et al. 2014), as was the case here. For this reason, including all possible site-pairs was not considered a good sampling choice. In addition, previous studies using GDM had found that model fit did not improve by increasing the number of site-pairs to more than 75,000 (Rosauer et al. 2014).

In some areas within the study region, there was very little data simply because little sampling had occurred in these areas. Therefore, there were many 0.1 degree grid cells with only one species. This could result in predictions of low similarity (high dissimilarity) especially if using the Sørensen metric (Kreft & Jetz 2010; Koleff et al. 2003). This would not be reflective of actual dissimilarity between sites but of differences in sampling effort across those sites. Only cells having five or more species were selected to train the model in order to control for this sampling effect (Kreft and Jetz 2010). There were 4,382 sites with five or more species (65% of grid cells with species data). From these, a random sample of 1,100 sites (25% coverage)

was selected to train the model. This represented 604,450 pair-wise comparisons. The average number of species per site in this sample was 17 species. **Appendix I** provides a summary of the species with the highest occurrence counts.

This analysis used all the sites from the sample. B_{sim} and B_{sor} dissimilarity values were calculated between all site-pairs in order to generate a site-pairs dissimilarity matrix for each measure. These site-pairs with their corresponding dissimilarity value represented the biological response for the GDM model (Rosauer et al. 2014). The matrix was reformatted into a table for input to the GD Modeller software. All sampling, generation of matrix, and reformatting was done in R v. 2.13.2 (R Development Core Team 2011).

2.2.4.3 Fitting the model

The model fitting process implemented with the GD Modeller software involved generating spline functions and coefficients to apply to the original predictor data to best fit the model. Through the fitting of I-spline functions, coefficients are derived that can be applied to the original raster predictor grids to best model the relationship between distance along the predictor gradient and compositional dissimilarity (the response variable) (Ferrier et al. 2007). The number of splines can influence the model fitting process, since too many splines would lead to over-fitting. In this analysis, the standard number of three splines was applied (Ferrier et al. 2007). New raster grids produced by transforming the original predictor grids can be used in GD Modeller to classify areas according to similarity in species composition. The GDM process was described in detail in Ferrier et al. (2007).

Initially, all 19 bioclimatic predictor grids, the DEM grid (elevation), and geographic distance were used to characterize climatic suitability at each of the selected 1,100 sites in the random sample. Actual evapotranspiration was also added as a predictor in some models (**Table 2.1**).

2.2.4.4 Backward elimination

A process of stepwise backward elimination was used to select the set of climatic variables that best predicted compositional dissimilarity based on ecological distance, as has been described in other studies (Rosauer et al. 2014; Williams et al. 2012). Initially, the model was fitted with all 21 predictor variables. By stepwise elimination, those predictors that did not contribute to the model were removed. Firstly, a single predictor that contributed least to the model was removed. Secondly, the model was fitted again with the remaining predictors. If the removed predictor caused a reduction of <0.01% in model deviance, it was dropped from the model (Rosauer et al. 2014). This process was repeated iteratively for all predictors individually, testing the reduction in model deviance as each predictor was removed. The final set of predictors holds those that could not be removed without causing a reduction in model deviance of >0.01% (Rosauer et al. 2014). An advantage of using this variable selection procedure is that it accounts for colinearity between predictors, since a strong predictor may still be removed if its relationship to one or more of the other predictors means that its removal causes little or no reduction in deviance explained (Rosauer et

al. 2014). Through this process, the set of model predictors was reduced to the most parsimonious set with minimal loss in model performance.

2.2.5 Model comparisons

Several different GDM models were performed with the B_{sor} and B_{sim} indices using the first random sample. These models were implemented using (i) different combinations of four variables (geographic distance, elevation, precipitation, temperature) and (ii) using all bioclimatic variables plus geographic distance and elevation (DEM). The deviances explained (DE) obtained from these models were compared in order to assess model performance.

Past studies explaining plant species richness patterns have been criticized for using elevation as a direct variable (Pausas and Austin 2001). It is argued that elevation is an indirect environmental variable that can be correlated to one or more direct variables related to plant or resource availability, but itself has no impact on these factors (Pausas and Austin 2001). In this case, elevation could be correlated to both temperature and precipitation, both of which are direct variables that impact on plant growth and resource use.

One variable that can be used as a measure of water use, but that is missing from the Worldclim bioclimatic data set is actual evapotranspiration (AET). In this analysis, I implemented some models using AET as a predictor variable, and removing the elevation predictor. The deviances explained (DE) obtained from these models were compared in order to assess model performance.

2.2.6 Model evaluation

Model validation is a crucial part of predictive modelling. In the case of species distribution modelling, where predictions may consist of the probabilities of a species' presence, models can be validated using presence/absence data and statistics such as the area under the receiver operator characteristic (ROC) curve (Pearce and Boyce 2006). In the case of GDM, however, such model evaluation proved problematic since predictions were based on similarities between sites rather than possible presence within a site. GDM model validation was also problematic because the data used to train the model was presence only data and so the model is subject to the same shortcoming as other models based on presence only data, namely, the absence of a true binary statistic (Pearce and Boyce 2006).

In order to evaluate model performance, a bootstrapping strategy (sampling with replacement) was adopted. Three random samples each containing 1,100 sites (see **Section 2.2.4.2**) were obtained from the entire data set. A species GDM model was performed for each sample using the most logical set of predictors available, and using the B_{sim} index (except for sample #1, for which both indices were used in order to compare the effect of using a different beta diversity index as explained in **Section 2.2.4.2**). The DE values obtained from

these models were compared as a way to evaluate model performance. The correlations between predicted and observed values obtained for each model were also used to evaluate model performance.

Summary tables for each data sample are shown in Appendix I.

2.2.7 Visualizing areas of most and least similarity

To visualize the spatial configuration of areas most and least similar to other areas across the region, the effective habitat area (EHA) for each grid cell was calculated and visualized in GIS maps. The EHA referred to the total area of similar habitat for each grid cell. In this case, the value for each cell represented the proportion of the total area of cells with a similar environment, and therefore similar biological composition, assuming that both cells were in a pristine condition (Williams et al. 2010). EHA is calculated from the focal cell *i* to all other *n* cells of interest. Since EHA is a proportion, obtained by multiplying the similarity of the focal cell *i* to cell *j* by the habitat condition of cell *j*, it is a unitless quantity. For any given cell, the effective area of similar habitat is the sum of the contributions for all *n* cells in the study region. The EHA of cell *i* (EHA_i) is:

$$EHA_i = \sum S_{ij}C_j \qquad \text{(equation 1)}$$

The EHA of each cell is calculated using a given "condition" layer *C*. In this case, the condition is one as it is the assumed pristine condition with MSA = 1 or all habitat intact, therefore, the EHA of cell *i* is simply the sum of its similarities ($\sum S_{ij}$) with all other cells. In this case, the range of EHA values (proportion of all cells that is similar to focal cell) calculated across all cells would vary between zero and 604,450. The highest boundary value would be the total number of piar-wise comparisons, i.e., 604,450 (see Methods 2.2.4.2). This would be the case if all cells were totally similar to each other, and all cells had pristine habitat (MSA=1) as it is assumed in this analysis. The lowest boundary value would be zero, in the case where all cells were totally dissimilar to each other, and all cells had MSA=1. A high value indicates that a cell has a high proportion of other cells similar in habitat to itself. A low value indicates that a cell has a low proportion of other cells similar to itself.

The EHA per grid cell was calculated for all three models and the results visualized in GIS maps. The results were compared visually among all three models as a way of evaluating whether the different samples produced different spatial patterns. A consistent spatial pattern produced by all three models would indicate that it was a robust representation of areas of low similarity or high similarity to other areas throughout the region. For sample #1, this was done for both the B_{sor} and B_{sim} models.

2.2.8 Classification by similarity

In order to visualize the spatial structure of species compositional turnover, the study area was classified and coloured, based on predicted dissimilarity (B_{sim}) from sample #1 model with highest DE, using the multi-

dimensional scaling method implemented in previous studies (Ferrier et al. 2007; Rosauer et al. 2014). All grid cells were clustered into 35 classes based on UPGMA hierarchical classification of predicted B_{sim}. Each class was coloured based on multi-dimensional scaling of predicted similarity, so that similar colours indicate compositional similarity (Rosauer et al. 2014). Grid cells were clustered into 35 classes in order to compare this classification based on compositional similarity with the WWF ecoregional classification of the study region (Olson et al. 2001).

2.3 RESULTS

2.3.1 Model fit varies with beta diversity index used

The deviance explained (DE) differed depending on the beta diversity index used. Total DE for the beta Sørensen (B_{sor}) model was 23.65% while the beta Simpson (B_{sim}) model had a DE of 14.97%. The predictor graphs in **Figure 2.1** and **Figure 2.2** show the spline functions for the final predictors in the B_{sor} and B_{sim} models, respectively. The B_{sor} model retained eight predictors and the B_{sim} model retained seven predictors after backward elimination. Both models had most predictors in common; however, the total amount of species turnover predicted by these predictors and hence their importance was different for each model (**Figure 2.1** & **Figure 2.2**). For example, in the B_{sor} model, the most important four predictors were geographic distance, precipitation of wettest quarter, maximum temperature of warmest month, and elevation. In the B_{sim} model, the most important four predictors were elevation, geographic distance, maximum temperature of warmest month, and precipitation seasonality. The B_{sor} model retained two predictors not retained in the B_{sim} model, namely, mean temperature of wettest quarter and precipitation of coldest quarter. Annual precipitation was retained as an important predictor in the B_{sim} model but not in the B_{sor} model.

The predictor plots or graphs (**Figure 2.1** & **Figure 2. 2**) show the shapes of the spline functions that best predicted species turnover (distance in species composition) based on the distance along each predictor retained by the model. For the B_{sor} model, for example, species turnover increased rapidly between distances of 0 to 25% along the geographic distance gradient (**Figure 2.1**). In this case, this means species turnover increased rapidly for sites that had a distance between them from 0 to 7 decimal degrees. Sites with distances between them of 25% to 50% along this gradient (7 to 13 decimal degrees) predicted a more gradual increase in species turnover. Differences between sites greater than 50% (13 to 26 decimal degrees) again predicted a steeper increase in species turnover. For the B_{sim} model, the predicted relationship between distance along the elevation gradient and species turnover is almost linear, suggesting that distance in elevation predicts a proportional increase in species turnover. For this model, precipitation seasonality and mean diurnal range also show a near linear relationship with species turnover (**Figure 2. 2**). For the B_{sor} model, the predictors that showed near linear relationships with species turnover were elevation and mean diurnal range (**Figure 2.1**). The other predictors in both models showed more curvilinear relationships with species turnover.

2.3.2 Spatial pattern of species compositional turnover

In general, the spatial pattern of effective habitat area (EHA) was similar for both B_{sor} and B_{sim} models. The values on the map scale indicate the range of EHA values (proportion of all cells that is similar to focal cell) calculated across all cells. The highest boundary value would be the total number of piar-wise comparisons, i.e., 604,450 and the lowest boundary value would be zero (see **Methods 2.2.7**). A high value indicates that a cell has a high proportion of other cells similar in habitat to itself. A low value indicates that a cell has a low proportion of other cells similar to itself. Both models showed that grid cells within ecoregions in Mexico, at the western limit of the study region, had low EHA (shown in red; **Figure 2.1** & **Figure 2.2**). This means that the similarity between these areas and other areas throughout the region is low. These areas can therefore be thought of as areas representing species assemblages not found elsewhere in the region. The other areas that showed low EHA corresponded to those areas in Costa Rica and Panama, and more specifically the Talamancan Montane Forests ecoregion, south of the study region. Areas within the central part of the study region that showed intermediate values of EHA (shown in yellow) included Central American Pine-Oak Forests, Central American Montane Forests, and Central American Dry Forests.

Subtle differences in spatial patterning were seen between the B_{sor} and B_{sim} models. In the B_{sor} model, areas of low EHA (shown in red) at the western limit of the study region encompassed the Southern Pacific Dry Forests ecoregion, located along the southern Pacific coast of Mexico. In the B_{sim} model, this ecoregion showed higher values of EHA. Likewise within Costa Rica, at the south end of the study region, the B_{sor} model showed grid cells within the Isthmian-Atlantic Moist Forests ecoregion as having lower EHA values (shown in orange; **Figure 2.1**) than they did in the B_{sim} model. In the B_{sim} model, the areas of low EHA in the south of the study region were confined to the Talamancan Montane Forests. Finally, within the central part of the study region and the Yucatan peninsula, areas of intermediate EHA (shown in yellow) were more widespread in the B_{sor} model than in the B_{sim} model.

The B_{sor} model showed lower values of EHA for all grid cells throughout the region. The lowest EHA value in the B_{sor} model was 43% lower than the lowest EHA value in the B_{sim} model. Likewise, the highest EHA value in the B_{sor} model was 28% lower than the highest value in the B_{sim} model. This difference in estimated EHA values was due to the way the specific indices calculated similarities between cells. These methodological differences are explained in the Discussion section.

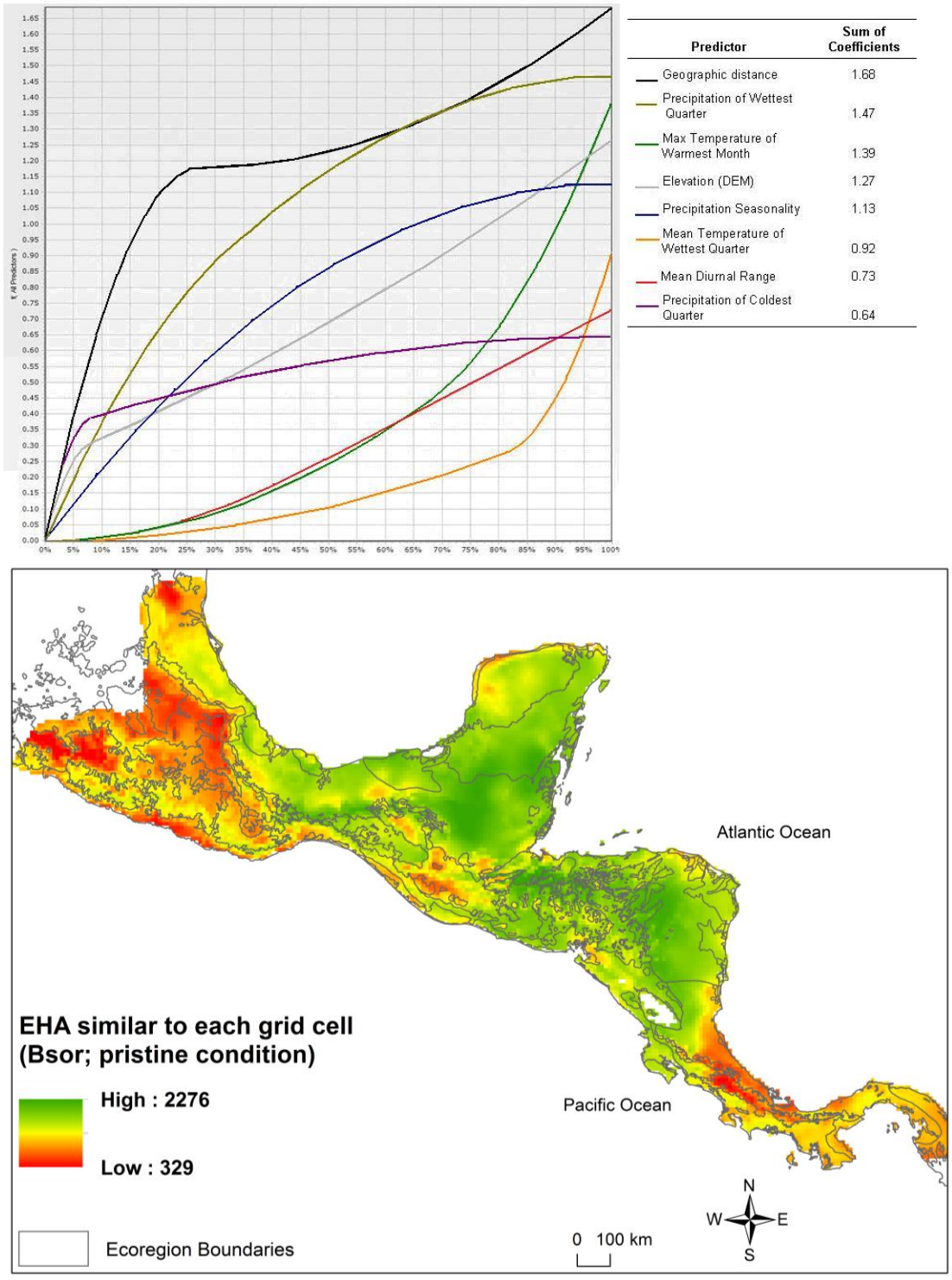


Figure 2.1 Predictor plot from B_{sor} model showing predictor spline functions fitted for each predictor (x-axis represents % change along the predictor gradient, the y-axis represents B_{sor} similarity). The sum of coefficients (height of the y-axis) on the plot represents the total amount of turnover predicted across the full range of sampled sites for this predictor. And, map showing spatial pattern of predicted EHA, the proportion of the total area of cells with similar habitat. A high value indicates that a cell has a high proportion of other cells similar in habitat to itself. A low value indicates that a cell has a low proportion of other cells similar to itself. EHA is a unitless quantity.

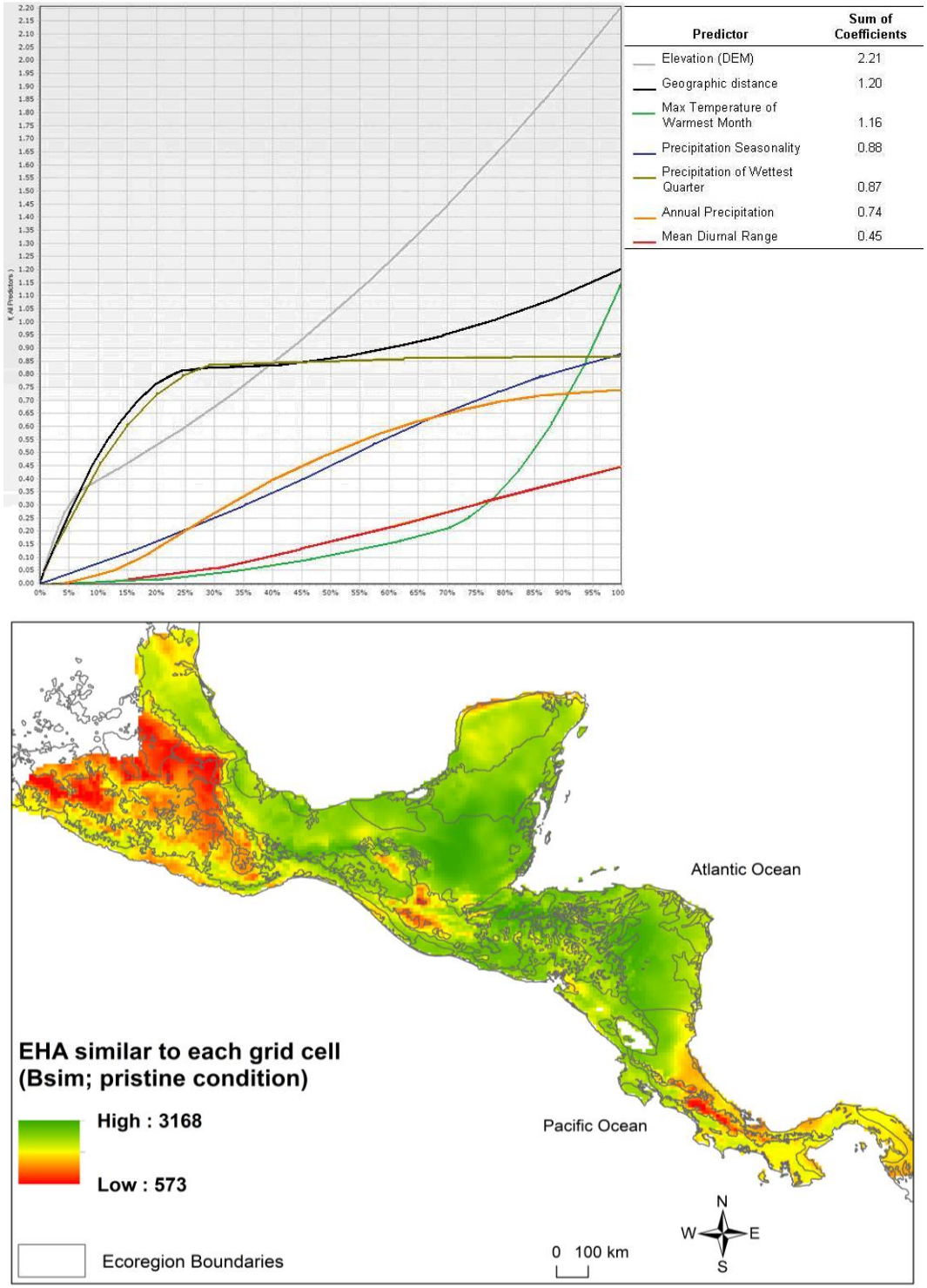


Figure 2.2 Predictor plot from B_{sim} model showing predictor spline functions fitted for each predictor (x-axis represents % change along the predictor gradient, the y-axis represents B_{sim} dissimilarity). The sum of coefficients represents the total amount of turnover predicted across the full range of sampled sites for this predictor. And, map showing spatial pattern of predicted EHA, the proportion of the total area of cells with similar habitat. A high value indicates that a cell has a high proportion of other cells similar in habitat to itself. A low value indicates that a cell has a low proportion of other cells similar to itself. EHA is a unitless quantity.

2.3.3 Predictors of compositional turnover

2.3.3.1 Models including elevation as a predictor

In addition to the models presented before, I also used different combinations of predictors to implement models using the B_{sor} and B_{sim} indices. The results from these models are presented in **Figures 2.3 - 2.5**. Only the predictor combinations that resulted in the highest DE are presented here (see **Appendix II** for total DE and sum of coefficients for models with other predictor combinations). The four-predictor model had the highest DE for both indices (**Figure 2.5**), followed by the three-predictor model and then the two-predictor model, as would be expected. The B_{sor} four-predictor model (DE=21.31%) showed a DE that was only 2.34% lower than the all-predictor model; similarly, the B_{sim} four-predictor model (DE=13.33%) had a DE only 1.64% lower than the B_{sim} all-predictor model. However, the rank or importance of each predictor differed between models. The B_{sor} models showed the most variation in predictor importance. Geographic distance was the most important predictor in the two-predictor model; however, when annual precipitation was added to the model, then elevation emerged as the most important predictor. When temperature was added in the four-predictor model, then annual precipitation emerged as the most important predictor. As shown in **Figure 2.1**, the all-predictor model again showed geographic distance as the most important, but elevation appeared as the fourth most important. In contrast, the B_{sim} models showed more consistency in predictor importance. Elevation was the most important predictor across all four models, including the all-predictor model (**Figures 2.2 - 2.5**). In the three and four-predictor models, elevation was followed by annual precipitation and geographic distance in importance (**Figure 2.4** and **Figure 2.5**). For the B_{sim} models, geographic distance was consistently in the top four predictors across all four models.

For the lone-predictor models (**Figure 2.6**), the geographic distance only model explained almost half as much of the variation in species composition (beta diversity) as the all-predictor models for both indices. The B_{sor} geographic distance model had a DE of 11.99% compared to the all-predictor model that had a DE of 23.65%. The B_{sim} geographic distance model showed a DE of 6.19% compared to the all-predictor model that showed a DE of 14.97%. For the B_{sor} index, the lone-predictor models using elevation, precipitation, and temperature separately showed similar values of DE that were above 7%. For the B_{sim} index, the lone-predictor model using elevation showed similar DE to geographic distance, while the models using temperature and precipitation separately showed respectively decreasing values of DE (**Figure 2.6)**.

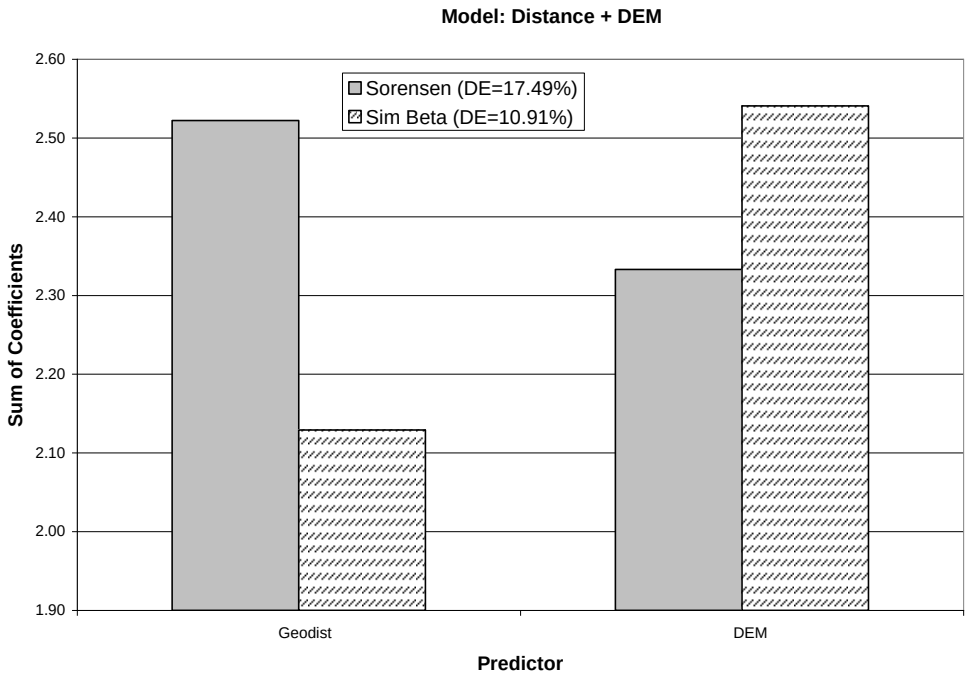


Figure 2.3 Graph showing total amount of turnover (sum of coefficients) predicted by geographic distance and elevation in the two-predictor model with the highest deviance explained (DE), for both beta diversity indices.

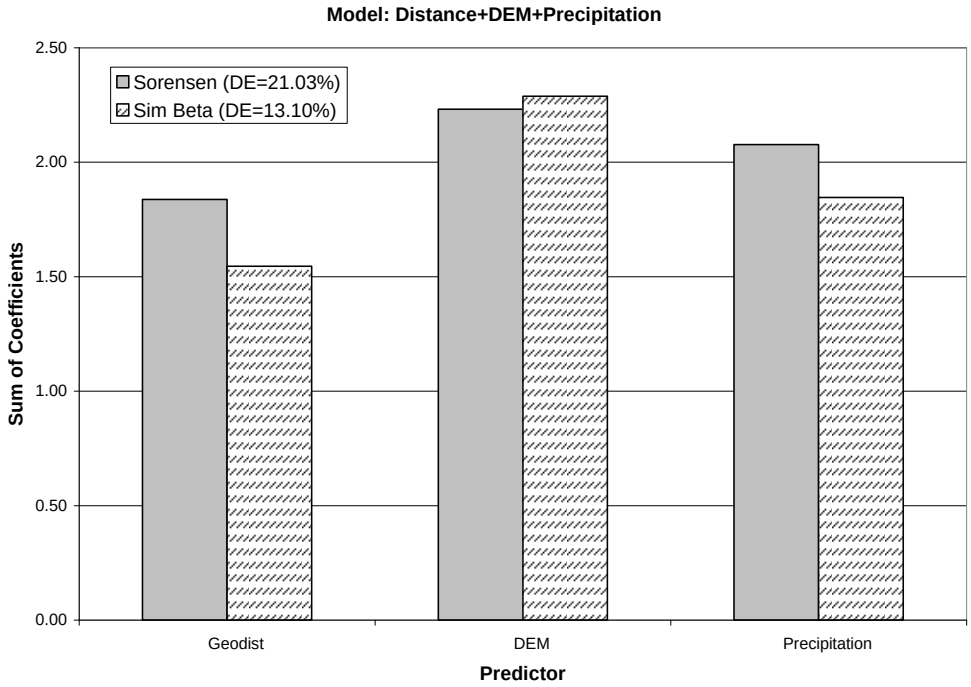


Figure 2.4 Graph showing total amount of turnover (sum of coefficients) predicted by different predictors in the three-predictor model with the highest DE, for both beta diversity indices.

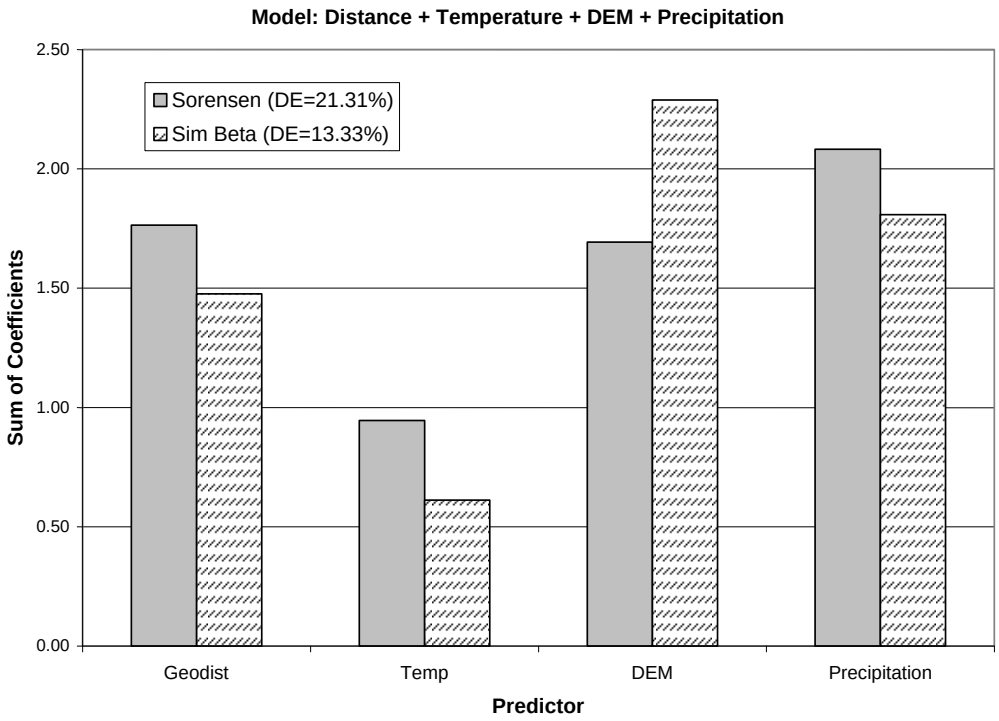


Figure 2.5 Graph showing total amount of turnover (sum of coefficients) predicted by different predictors in the four-predictor model with the highest DE, for both beta diversity indices.

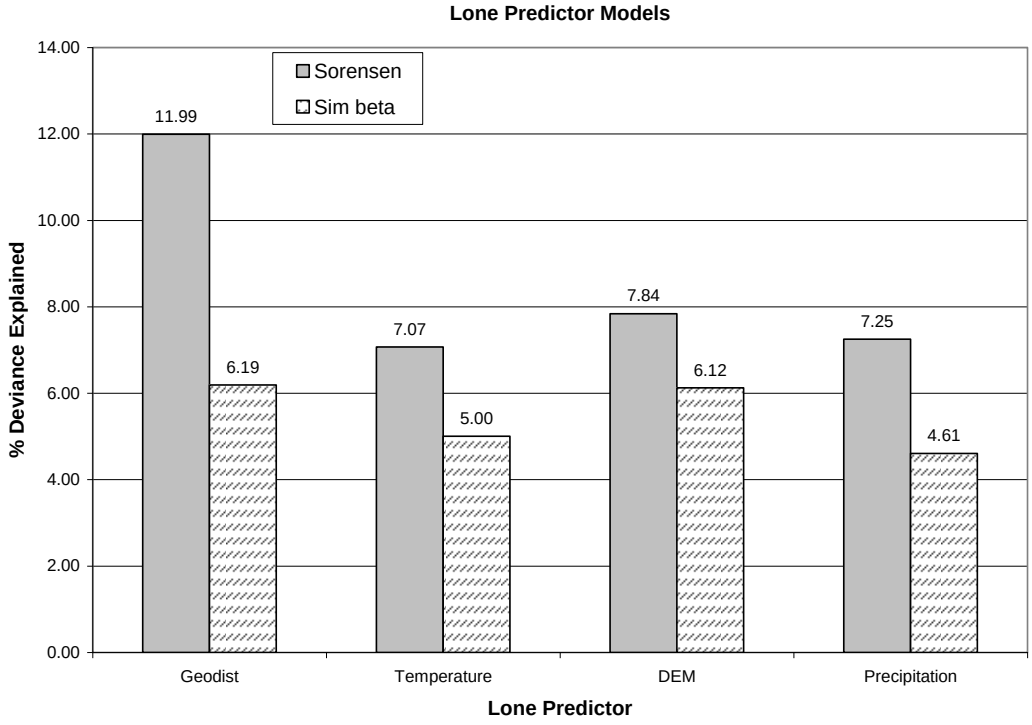


Figure 2.6 Graph showing DE for each lone-predictor model, for both beta diversity indices.

2.3.3.2 Models including actual evapotranspiration (AET) as predictor

Since the B_{sim} models showed more consistency in predictor importance, subsequent models were performed using this index. In addition, the elevation variable was removed, and the AET variable was added to the set of model predictors as indicated previously (see **Section 2.2.4.3**). The results of the GDM process performed on the three separate data samples are shown in **Figures 2.7 - 2.9**. The model from sample #1 had the highest DE of 14.92%, followed by sample #2 model with DE of 14.22%, and finally sample #3 model with DE of 13.81%.

The three GDM models each reduced to a different number of predictors. The model from sample #1 retained eleven predictors, sample #3 retained nine predictors, and sample #2 had the least at seven predictors retained. All three models showed six of the same predictors in common (mean temperature of wettest quarter, geographic distance, precipitation of wettest quarter, precipitation seasonality, mean diurnal range, minimum temperature of coldest month); however, the importance of these predictors differed among models (**Figures 2.7- 2.9**). For all three models, mean temperature of wettest quarter and geographic distance were both among the top three predictors. For sample #1 and sample #2, precipitation of wettest quarter was also among the top three predictors (**Figure 2.7** & **Figure 2.8**). For sample #3, mean diurnal range was the second most important predictor (**Figure 2.9**). In the model from sample #1, precipitation of coldest quarter and mean temperature of driest quarter were retained as predictors but not in the other models. In the model from sample #2, precipitation of driest quarter was retained but not in the other models. Actual evapotranspiration was retained in the sample #1 and sample #3 models.

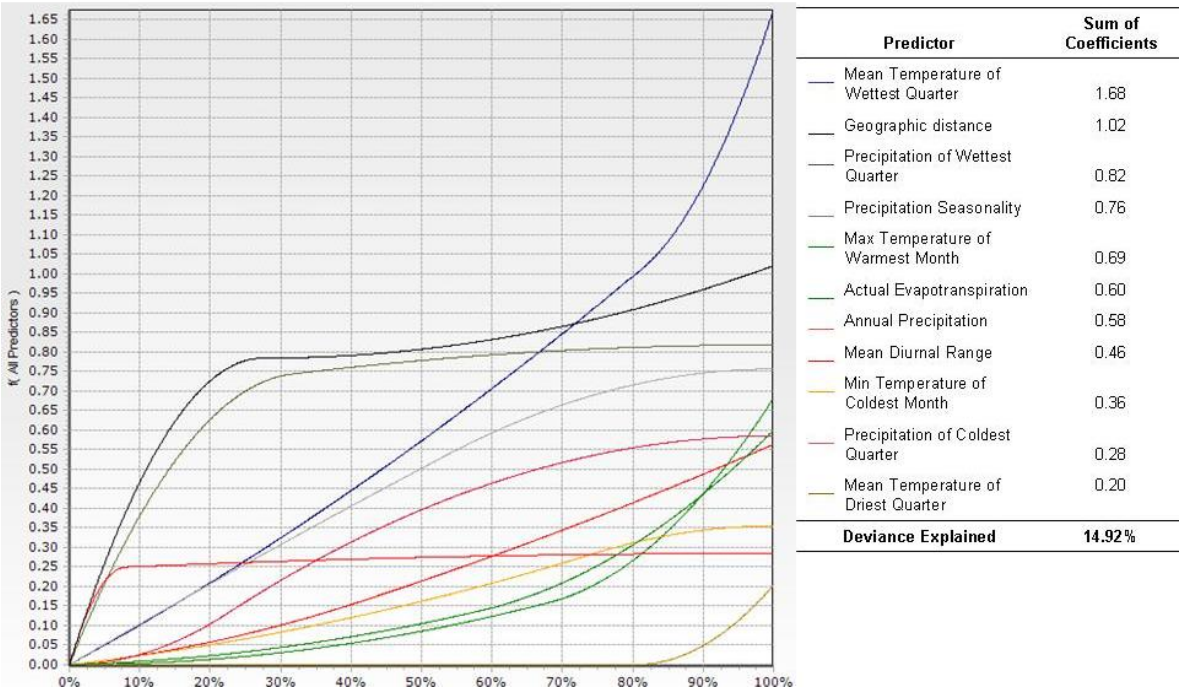
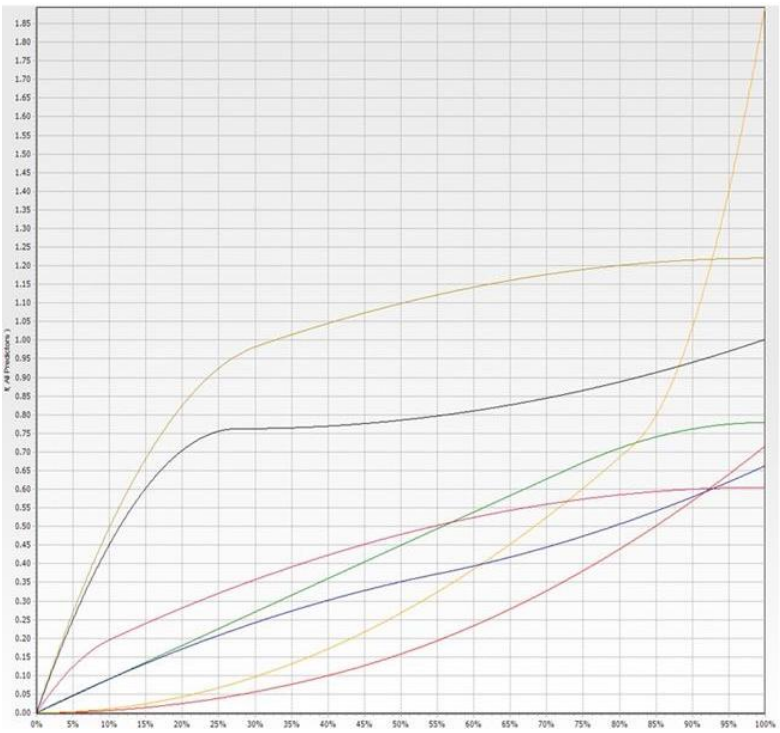
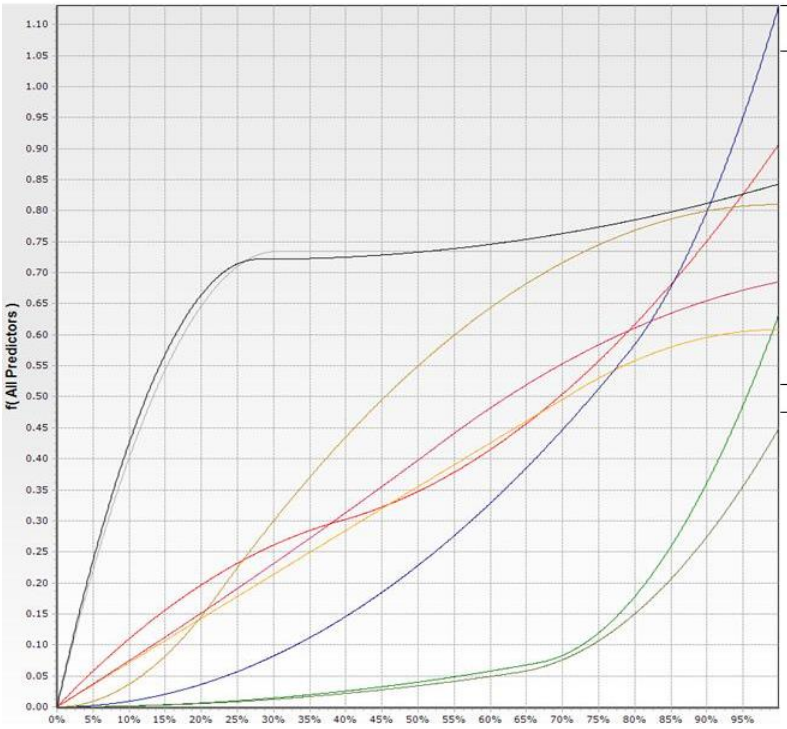


Figure 2.7 Predictor plot for sample #1 B_{sim} model showing predictor spline functions fitted for each predictor. The x-axis represents % change along the predictor gradient and the y-axis represents B_{sim} dissimilarity.



Predictor	Sum of Coefficients
Mean Temperature of Wettest Quarter	1.90
Precipitation of Wettest Quarter	1.22
Geographic distance	1.00
Min Temperature of Coldest Month	0.78
Mean Diurnal Range	0.72
Precipitation Seasonality	0.66
Precipitation of Driest Quarter	0.61
Deviance Explained	**14.22%**

Figure 2.8 Predictor plot for sample #2 B_{sim} model showing predictor spline functions fitted for each predictor. The x-axis represents % change along the predictor gradient and the y-axis represents B_{sim} dissimilarity.



Predictor	Sum of Coefficients
Mean Temperature of Wettest Quarter	1.13
Mean Diurnal Range	0.91
Geographic distance	0.84
Annual Precipitation	0.81
Precipitation of Wettest Quarter	0.74
Precipitation Seasonality	0.69
Max Temperature of Warmest Month	0.63
Min Temperature of Coldest Month	0.61
Actual Evapotranspiration	0.45
Deviance Explained	**13.81%**

Figure 2.9 Predictor plot for sample #3 B_{sim} model showing predictor spline functions fitted for each predictor. The x-axis represents % change along the predictor gradient and the y-axis represents B_{sim} dissimilarity.

2.3.4 Model evaluation

2.3.4.1 Low model fit and similar spatial pattern of EHA

The results showed that model performance was dependent on the data sample that was used, since DE values varied among the three models (**Figures 2.7 - 2.9)**. However, the spatial pattern of predicted EHA based on the input predictor variables showed robustness across all three models. The mean DE across all three models was 14.32% (S.D.= 0.56). The low DE values for all models indicated low model fit and suggested that the ability of the models to predict variation in species composition was low.

Overall, the spatial pattern of effective habitat area (EHA) was similar for models based on the three data samples. The range of EHA values differed among the models and sample #2 showed the widest range of values. All models showed that grid cells having the lowest EHA (shown in red) were confined to ecoregions at the western limit of the study region (within Mexico) and to the Talamancan Montane Forests ecoregion at the south of the study area (**Figure 2.10**). Areas within the central part of the study region that showed intermediate (shown in yellow) to low values of EHA included Central American Pine-Oak Forests, Central American Montane Forests, and Central American Dry Forests. Ecoregions in Costa Rica and Panama that showed intermediate values of EHA included Isthmian-Atlantic Moist Forests, Panamanian Dry Forests, Eastern Panamanian Montane Forests, and Choco-Darien Moist Forests.

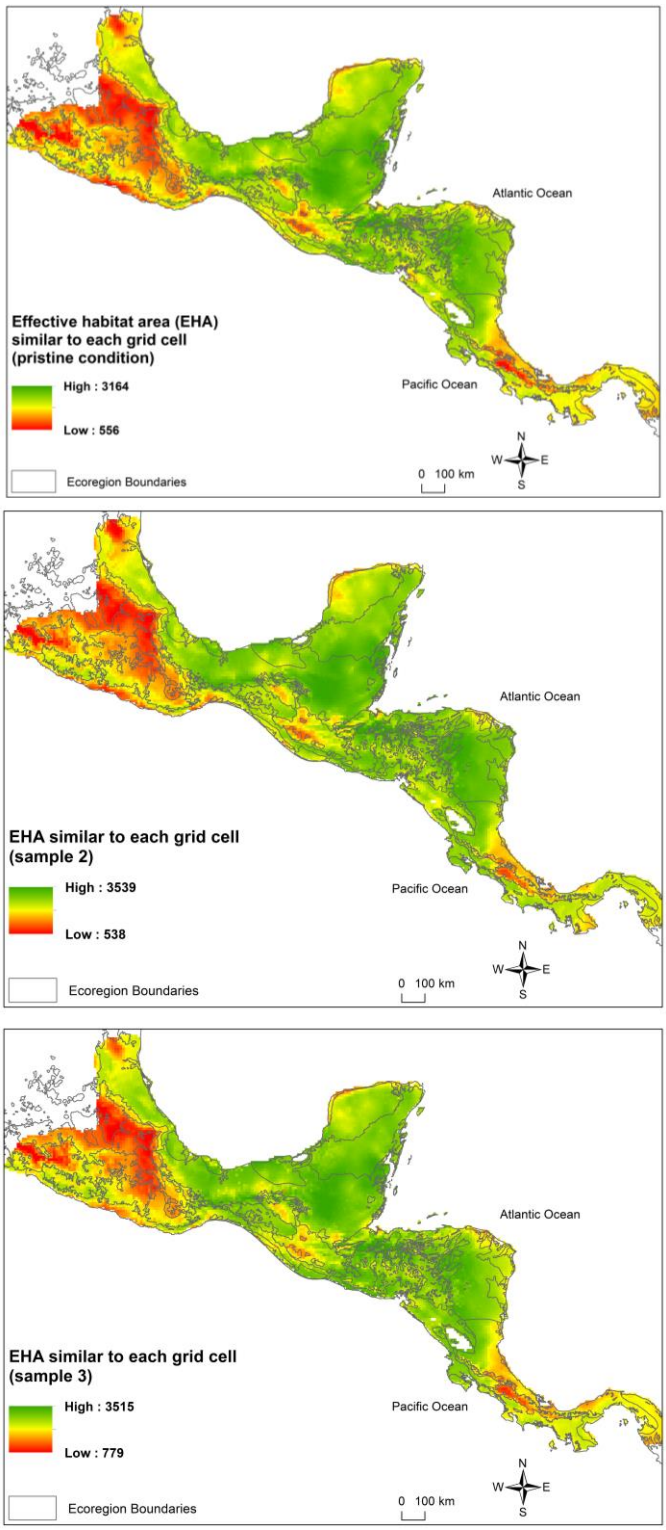


Figure 2.10 Maps showing predicted EHA (proportion of total area of cells with similar habitat) at pristine conditions for each grid cell, in all three samples of data. A high EHA value indicates that a cell has a high proportion of other cells similar in habitat to itself. A low EHA value indicates that a cell has a low proportion of other cells similar to itself. EHA is a unitless quantity.

2.3.4.2 Predictive performance

All the sample models showed low correlation between observed and predicted dissimilarity values (**Figure 2.11**). The intercept value for all models was near 0.8. This suggested that sites with minimal geographic distance and identical environmental conditions had a high dissimilarity value of 0.8. Any differences in geographic distance and environment only predicted changes in dissimilarity from 0.8 to 1. The low model fit and high dissimilarity values predicted from minimal geographic and environmental distance was a surprising result and suggested that the models were overpredicting dissimilarity (in this case B_{sim}). Examination of the model residuals also indicated that overprediction was occurring. Residuals were highly skewed (**Figure 2.11**). Most residuals were negative (below the zero line) suggesting that predictions were too high. Some of the residuals showed very high overestimates as they fell between -6 (sample #2) to -10 (sample #1).

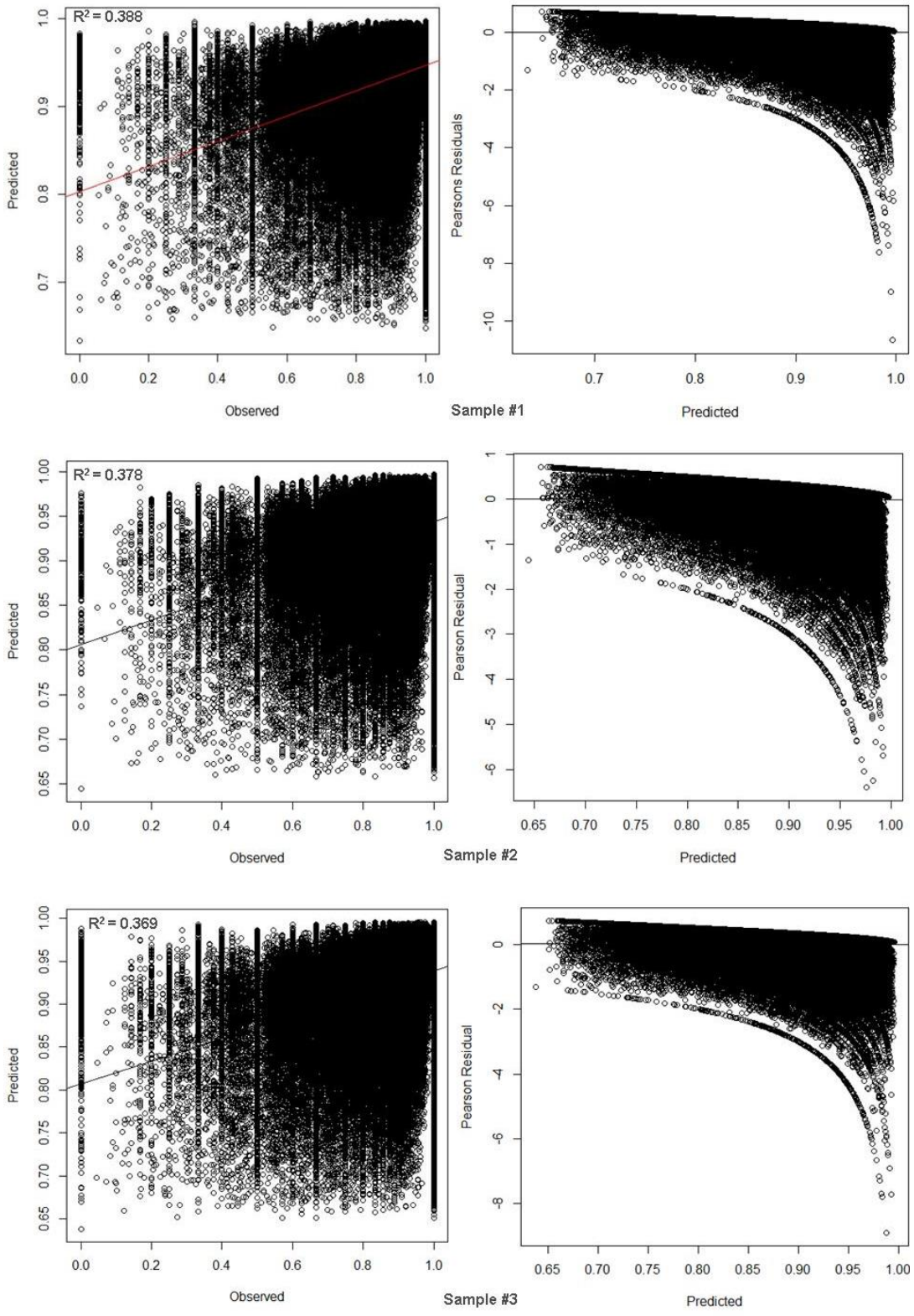


Figure 2.11 Correlation plots for observed and predicted pair-wise dissimilarity values (left panel) and residual plots (right panel) for models based on three separate data samples.

2.3.5 Classification by predicted similarity in tree species composition

Classification according to compositional affinity based on the eleven-predictor model results showed some broad similarities with the WWF ecoregional classification (Olson et al. 2001) (**Figure 2.12**). Broadly similar patterns were distinguished, such as the differentiation of three classes in the southern region of the study area: one class was on the Atlantic side of Costa Rica, one class was on its Pacific side, and one class separated these two (**Figure 2.12**). These classes were broadly congruent with Isthmian-Atlantic Moist Forests, Isthmian-Pacific Moist Forests, and Talamancan Montane Forests respectively. Within Honduras and south of the Yucatan peninsula (which includes the Chiapas region), the different classes could be interpreted as broadly congruent with the complex of WWF ecoregions represented there: Central American Pine-Oak Forests, Central American Dry Forests, Central American Montane Forests, Chiapas Depression Dry Forest, Chiapas Montane Forest and Sierra Madre de Chiapas Moist Forests (**Figure 2.12**).

A major difference between the two classifications was that the B_{sim} classification showed more classes in the Mexican area and the Yucatan area of the study region than in the south (**Figure 2.13, A-C**). In the Mexican area, the B_{sim} classification showed more classes than identified in the same region by the WWF ecoregional classification (**Figure 2.13, A**). Another difference in this region was that the B_{sim} classification showed only one class for the WWF-defined Veracruz Montane Forests and Oaxacan Montane Forests (**Figure 2.13, A, see arrow**). In addition, the B_{sim} classification did not distinguish between Yucatan Moist and Yucatan Dry Forests (**Figure 2.13, B**). This classification also showed three separate classes within the WWF-defined Peten-Veracruz Moist Forests (**Figure 2.13, B, see arrow**). In the south, the B_{sim} classification did not distinguish classes which occured in the WWF ecoregional classification, such as Costa Rican Seasonal Moist Forests, Panamanian Dry Forests, Eastern Panamanian Montane Forests, Choco-Darien Moist Forests, and Southern Mesoamerican Pacific Mangroves (**Figure 2.13, C**).

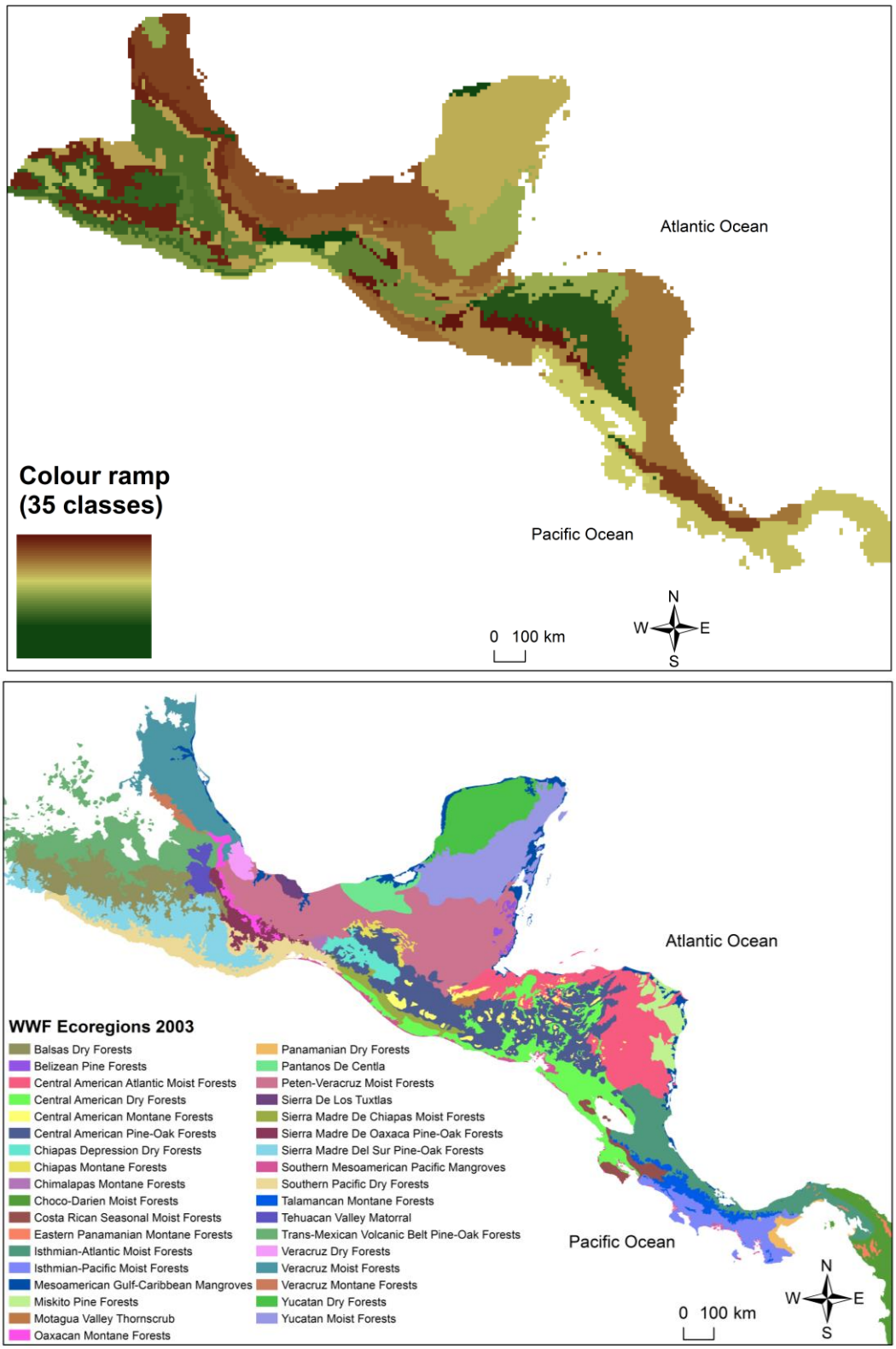
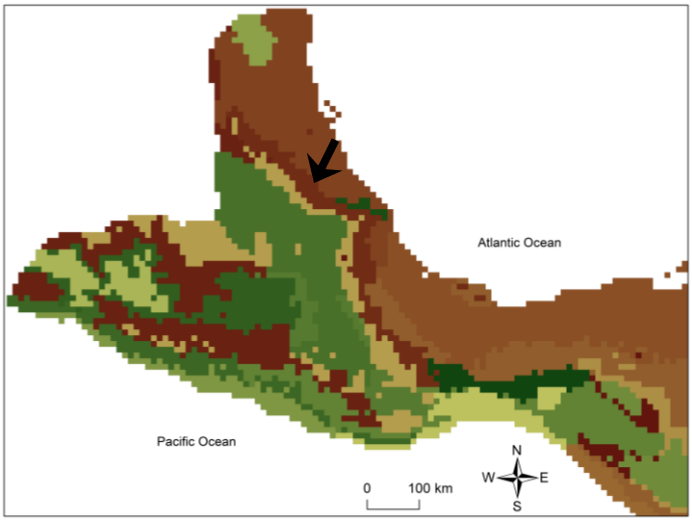
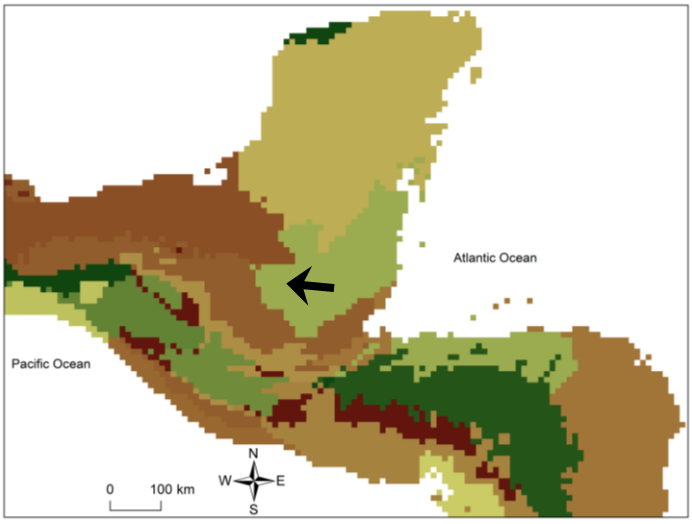


Figure 2.12 Classification by predicted species similarity based on the eleven-predictor model using the B_{sim} index (top). Different colour classes indicate groups of grid cells that are similar in species composition. Number of classes was restricted to 35 for comparison with WWF 2003 ecoregional classification (bottom).

A



B



C

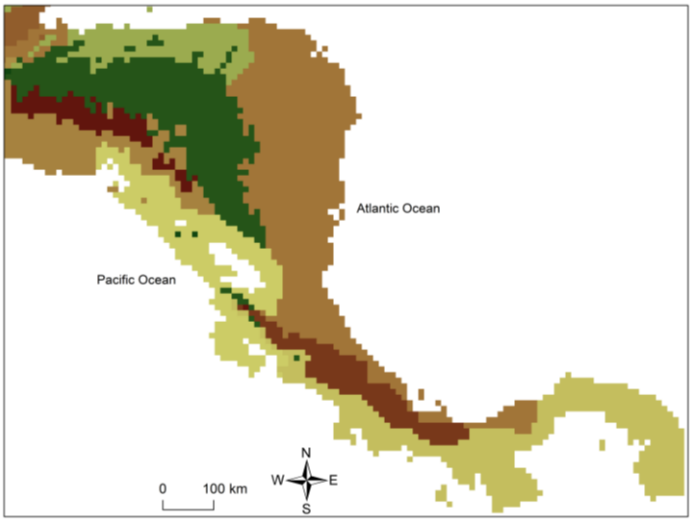


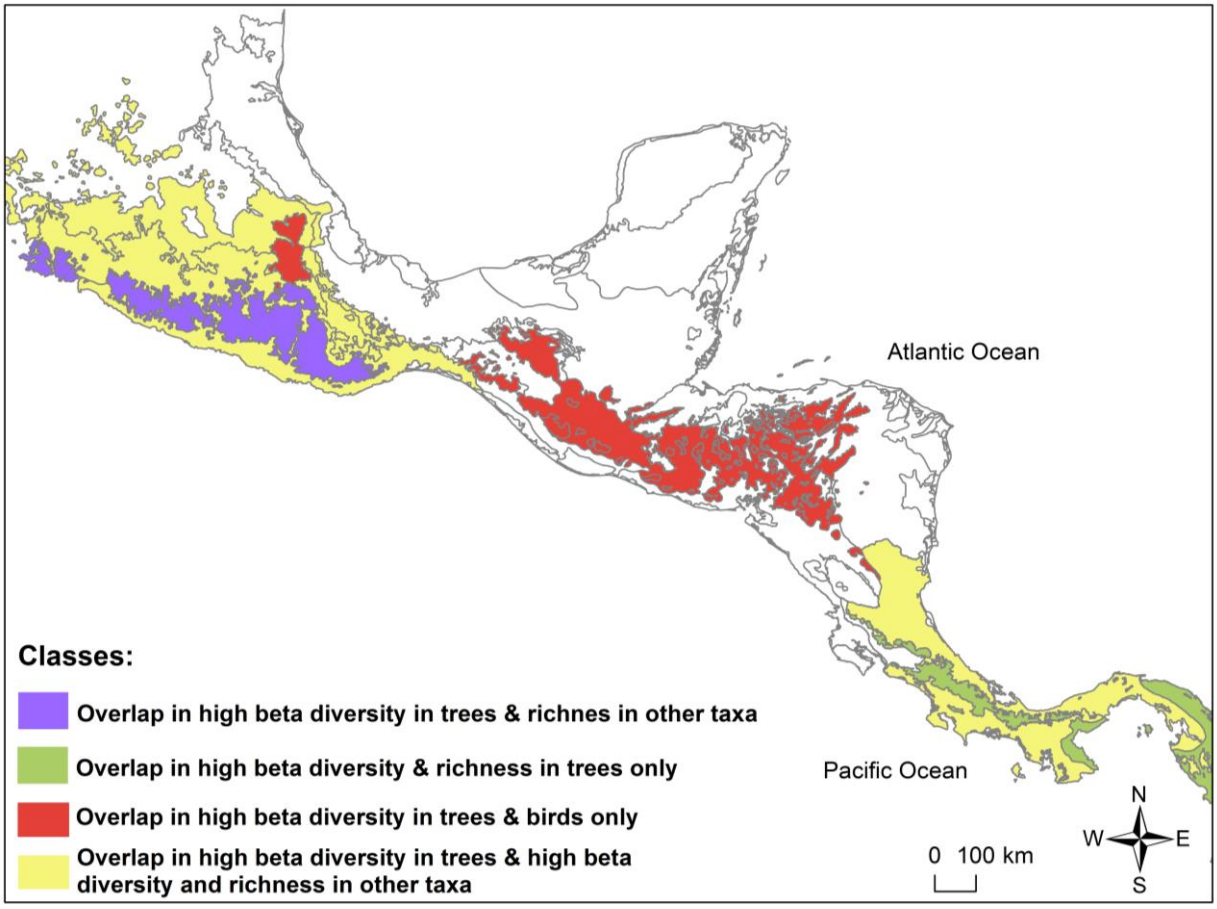
Figure 2.13 Zoomed-in views for different regions of the study area: A) Mexico B) Yucatan and C) South.

2.3.6 Spatial congruence in patterns of beta diversity and patterns of species richness

The pattern of predicted EHA presented here showed clusters of low EHA (high beta diversity) throughout the region, surrounded by large areas of high EHA (low beta diversity) (**Figure 2.10**). The results of this chapter suggest both congruence and decoupling of patterns of tree beta diversity, as exemplified by EHA, when compared to patterns of species richness or beta diversity observed in other studies (**Figure 2.14**).

Previous studies conducted globally and in the Mesoamerican region using trees and other taxa have focused only on patterns of species richness (e.g. Barthlott et al. 2007). Models of potential species richness (PSR) showed a clear increasing gradient in PSR from north to south along the region (Golicher et al. 2012). In contrast, the pattern of predicted EHA presented here was less straightforward, and showed there were clusters of low EHA throughout the region. There was no clear north to south gradient in the predicted EHA pattern as with PSR. Nevertheless, there were several areas in which EHA and PSR patterns coincided. For example, Sierra Madre de Oaxaca Pine-Oak Forests and Oaxaca Montane Forests showed both high PSR and intermediate to low EHA (high beta diversity). PSR was highest in Costa Rica and Panama, and Talamancan Montane Forests had highest PSR (Golicher et al. 2012). This pattern is the same for EHA, since the model results presented here also showed the ecoregions to the south of the region as having lowest EHA (high beta diversity). In contrast, most of the ecoregions within Mexico (in the west of the study region) that this study identified as having low EHA (high beta diversity) showed low to intermediate PSR according to Golicher et al. (2012). Also, that study showed low PSR in Yucatan Dry Forests, but this ecoregion showed intermediate EHA in this study, indicating higher compositional turnover compared to many other areas in the region. Another study on global patterns of plant diversity found highest species richness in Balsas Dry Forests, Trans-Mexican Volcanic Belt Pine-Oak Forests and Choco-Darien Moist Forests (Kier et al. 2005). The present study also found these ecoregions to have low EHA or high beta diversity.

At the regional level, Anderson et al. (2008) showed areas of high species richness for birds, mammals and amphibians that corresponded to Sierra Madre de Oaxaca Pine-Oak Forests, Oaxaca Montane Forests, and Southern Pacific Dry Forests. In the present study, these areas showed intermediate to low EHA or high beta diversity, showing congruence in patterns of species richness and beta diversity within these ecoregions. Both studies also showed similar congruence for Isthmian-Atlantic Moist Forests, Isthmian-Pacific Moist Forests, Choco-Darien Moist Forests, and Eastern Panamanian Montane Forests. In contrast, in the Yucatan area of the study region, Anderson et al. (2008) showed Peten-Veracruz Moist Forests had high species richness, whereas in this study that ecoregion showed low beta diversity (high EHA). Similarly, they highlight Central American Atlantic Moist Forests as high in species richness, but this study showed it as showing low beta diversity (high EHA). Talamancan Montane Forests was identified as poor in species richness (Anderson et al. 2008) but high in beta diversity (low EHA). Regionally and globally, other studies have identified patterns of beta diversity for different taxa. Some ecoregions showed overlap in high beta diversity (low EHA) for trees and high beta diversity for birds (McKnight et al. 2007; Melo et al. 2009) and amphibians (Ochoa-Ochoa et al. 2012) (**Figure 2.14**). .



Ecoregion names colour coded by class (data from WWF Ecoregions 2003)

- Ecoregion boundaries
- Balsas Dry Forests
- Central American Montane Forests
- Central American Pine-Oak Forests
- Choco-Darien Moist Forests
- Eastern Panamanian Montane Forests
- Isthmian-Atlantic Moist Forests
- Isthmian-Pacific Moist Forests
- Oaxacan Montane Forests
- Panamanian Dry Forests
- Sierra Madre De Oaxaca Pine-Oak Forests
- Sierra Madre Del Sur Pine-Oak Forests
- Southern Pacific Dry Forests
- Talamancan Montane Forests
- Tehuacan Valley Matorral
- Trans-Mexican Volcanic Belt Pine-Oak Forests

Figure 2.14 Geographic variation in beta diversity and richness congruence of trees and other taxa, within Mesoamerica. White indicates no overlap in the focal taxa. Ecoregion names are colour-coded by the class they belong to.

2.4 DISCUSSION

The method presented here for modelling spatial patterns of species turnover can provide important information for conservation planning efforts within the Mesoamerican region. The method is based on Effective Habitat Area (EHA), or habitat remaining for originally occurring species, and is a measure that takes into account pair-wise similarity between sites (in this case between 1^0grid cells). The process

presented in this chapter identified areas of low and high EHA. Areas of low EHA represented areas that had overall low similarity to other areas in the region, indicating high beta diversity. They can therefore be thought of as areas with species assemblages not found elsewhere in the region. To my knowledge, this is the first study in the Mesoamerican region that has implemented the GDM process based on modelling beta diversity to identify areas that may warrant conservation efforts.

2.4.1 Spatial Pattern of Species Compositional turnover

2.4.1.1 Congruence in patterns of beta diversity and species richness

The results of this chapter provide evidence for congruence in patterns of tree beta diversity when compared to patterns of tree and vertebrate richness within the Mesoamerican region. Similarly, some ecoregions showed congruence in high beta diversity for trees and vertebrate taxa. This congruence maybe due to common histories of speciation and extinction (McKnight et al. 2007). Those authors showed patterns of congruence were especially apparent in Neotropical mountains. The present results also show high beta diversity in mountainous regions and help to highlight these areas as centres of diversity.

2.4.1.2 Classification by compositional similarity

The similarity between the classification presented in these results and the WWF ecoregions classification suggests that both of these approaches are capturing broad scale patterns of diversity driven by broad scale factors such as climate (Field et al. 2009). The B_{sim} classification highlighted some areas of differentiation, such as within the Peten-Vercruz Moist Forests ecoregion (**Figure 2.13B**) that may warrant further investigation e.g. with further collection surveys. Some of these demarcations may be due to differences in precipitation being captured by the model e.g. in Belize the southern region receives more cumulative rainfall than the northern region (**Figure 2.13B**). Further areas of differentiation present in the WWF ecoregions but not in this B_{sim} classification may highlight the importance of expert knowledge in refining broad scale differentiations.

2.4.2 Predictors of species compositional turnover

Geographic distance was consistently one of the most important predictors across all the models that were implemented, in concordance with other studies (Keil et al. 2012). When the models included geographic distance and elevation (DEM), the B_{sor} model showed the most important predictor was geographic distance. For the B_{sim} model, this was reversed and elevation was the most important predictor. When the elevation variable was removed, the B_{sim} model showed temperature variables as most important followed by precipitation or geographic distance. It is possible that elevation was masking the effect of temperature as a proximal predictor (Pausas and Austin 2001) and when it was removed the effect of temperature became more evident. This result could reflect the variation in plant adaptations to temperature and in particular adaptations to the narrower climatic niches of tropical mountains (Buckley and Jetz 2008), hence making temperature a more important predictor for species turnover patterns in plants.

Studies have documented the relationship between productivity and beta diversity (Field et al. 2009; He and Zhang 2009). In this study, productivity was not used as a predictor since it was not considered to be a proximal predictor related to a causal mechanism of plant diversity (Pausas and Austin 2001). This study focused on climatic predictors since climate has been documented as a determinant of both species richness and beta diversity at continental and global scales (Field et al. 2009; Svenning et al. 2010). Kiel *et al* (2012) found that the effect of climate was scale-dependent and it was more important at coarse grains. In addition, variables representing energy and water availability, such as those used here, have also been shown to be strong predictors of broad scale richness patterns in both plants and vertebrates (Chown et al. 2004).

The fact that distance was consistently among the top predictors across all models implemented is in accordance with other studies that have found a significant effect of distance in explaining variations in species diversity. Distance was found to account for 46.1% of variation in plant species richness across biomes (Gerstner et al. 2014). Similarly, a study on European mammals found that variation in species composition (beta diversity) accounted for purely by space was even greater than the portion attributed solely to environment (Svenning et al. 2010). This strong effect of distance suggests that species do not necessarily occupy all suitable habitat (Keil et al. 2012). This has important implications for the presumption that species engage in climate tracking. This means that other factors affecting the location of species must be taken into account in planning for climate change effects.

2.5 CONCLUSION

Broad scale analyses of beta diversity patterns are important tools for conservation planning. The use of beta diversity indices to detect spatial turnover patterns in trees can help to identify regions where many closely spaced protected areas may be needed to effectively protect diversity (McKnight et al. 2007). This is clearly the case with the ecoregions in Mexico highlighted in this chapter as having low EHA, and showing congruent patterns of high beta diversity in other taxa. Other areas may require a different spatial configuration, but it is important to remember that the planning of protected areas must also consider the human landscape and patterns of land use.

Once a GDM species model has been generated, it can be applied to answer questions relating to reserve gap analyses or to the effect of anthropogenic changes on species diversity. I specifically apply the model developed in this chapter to such questions in subsequent chapters.

2.6 REFERENCES

Allnutt, T. F., Ferrier, S., Manion, G., Powell, G. V. N., Ricketts, T. H., Fisher, B. L., Harper, G. J., Irwin, M. E., Kremen, C., Labat, J.-N., Lees, D. C., Pearce, T. a. and Rakotondrainibe, F., 2008. A method for quantifying biodiversity loss and its application to a 50-year record of deforestation across Madagascar. *Conservation Letters*, *1* (4), 173–181.

Anderson, E. R., Cherrington, E. A., Tremblay-Boyer, L., Flores, A. I. and Sempris, E., 2008. Identifying Critical Areas for Conservation : Biodiversity and Climate Change in Central America , Mexico , and the Dominican Republic. *Biodiversity*, *9* (3 & 4), 89–99.

Arponen, A., Moilanen, A. and Ferrier, S., 2008. A successful community-level strategy for conservation prioritization. *Journal of Applied Ecology*, *45* (5), 1436–1445.

Barthlott, W., Hostert, A., Kier, G., Kuper, W., Kreft, H., Mutke, J., Daud Rafiqpoor, M. and Henning Sommer, J., 2007. Geographic patterns of vascular plant diversity at continental to global scales. *Erdkunde*, *61* (4), 305–315.

Benito, B. M., Cayuela, L. and Albuquerque, F. S., 2013. The impact of modelling choices in the predictive performance of richness maps derived from species-distribution models: guidelines to build better diversity models. *Methods in Ecology and Evolution*, *4* (4), 327–335.

Buckley, L. B. and Jetz, W., 2008. Linking global turnover of species and environments. *Proceedings of the National Academy of Sciences of the United States of America*, *105* (46), 17836–41.

Cardoso, P., Borges, P. a. V. and Veech, J. a., 2009. Testing the performance of beta diversity measures based on incidence data: the robustness to undersampling. *Diversity and Distributions*, *15* (6), 1081–1090.

Cayuela, L., Gálvez-Bravo, L., Perez, R., de Alburquerque, F. S., Golicher, D. J., Zahawi, R. A., Ramírez-Marcial, N., Garibaldi, C., Field, R., Benayas, J. M. R., González-Espinosa, M., Balvanera, P., Castillo, M. Á., Figueroa-Rangel, B. L., Griffith, D. M., Islebe, G. A., Kelly, D. L., Olvera-Vargas, M., Schnitzer, S. A., Velázquez, E., Williams-Linera, G., Brewer, S. W., Camacho-Cruz, A., Coronado, I., Jong, B. de, Castillo, R. del, Cerda, Í. G. la, Fernández, J., Fonseca, W., Galindo-Jaimes, L., Gillespie, T. W., González-Rivas, B., Gordon, J. E., Hurtado, J., Linares, J., Letcher, S. G., Mangan, S. A., Meave, J. A., Méndez, E. V., Meza, V., Gaona, S. O.-, Peterson, C. J., Ruiz-Gutierrez, V., Snarr, K. A., Dzul, F. T., Valdez-Hernández, M., Viergever, K. M., White, D. A., Williams, J. N., Bonet, F. J. and Zamora, R., 2012a. The Tree Biodiversity Network (BIOTREE-NET): prospects for biodiversity research and conservation in the Neotropics. *Biodiversity & Ecology*, *4* , 211–224.

Cayuela, L., Granzow-de la Cerda, Í., Albuquerque, F. S. and Golicher, D. J., 2012b. Taxonstand: An r package for species names standardisation in vegetation databases. *Methods in Ecology and Evolution*, *3* (6), 1078–1083.

Chase, M. W. and Reveal, J. L., 2009. A phylogenetic classification of the land plants to accompany APG III. *Botanical Journal of the Linnean Society*, *161* , 122–127.

Chown, S. L., Sinclair, B. J., Leinaas, H. P. and Gaston, K. J., 2004. Hemispheric asymmetries in biodiversity--a serious matter for ecology. *PLoS Biology*, *2* (11), e406.

Colwell, R. and Coddington, J., 1994. Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society of Biological Sciences*, *345* (1311), 101–118.

Conservation International, 2004. *CEPF Ecosystem Profile: Northern Region Of The Mesomaerica Biodiversity Hotspot*. *Analysis*. Washington, D.C.

Cowling, R. M. and Pressey, R. L., 2001. Rapid plant diversification: planning for an evolutionary future. *Proceedings of the National Academy of Sciences of the United States of America*, *98* (10), 5452–5457.

Dunlop, M., Hilbert, D., Ferrier, S., House, A., Liedloff, A., Prober, S. M., Smyth, A., Martin, T. G., Harwood, T., Williams, K. J., Fletcher, C. and Murphy, H., 2012. *The implications of climate change for biodiversity conservation and the National Reserve System: Final Synthesis. A report prepared for the Department of Sustainability, Environment, Water, Population and Communities, and the Department of Climate Change*. Canberra.

Elith, J. and Leathwick, J., 2007. Predicting species distributions from museum and herbarium records using multiresponse models fitted with multivariate adaptive regression splines. *Diversity and Distributions*, *13* (3), 265–275.

Fairbanks, D., Reyers, B. and van Jaarsveld, A. S., 2001. Species and environment representation: selecting reserves for the retention of avian diversity in KwaZulu-Natal, South Africa. *Biological Conservation*, *98* (3), 365–379.

Faith, D., Ferrier, S. and Walker, P., 2004. The ED strategy: how species‐level surrogates indicate general biodiversity patterns through an "environmental diversity" perspective. *Journal of Biogeography*, *31* , 1207–1217.

Faith, D. P., Carter, G., Cassis, G., Ferrier, S. and Wilkie, L., 2003. Complementarity, biodiversity viability analysis, and policy-based algorithms for conservation. *Environmental Science & Policy*, *6* (3), 311–328.

Faith, D. P., Minchin, P. R. and Belbin, L., 1987. Compositional Dissimilarity as a Robust Measure of Ecological Distance. *Vegetatio*, *69* (1), 57–68.

Feeley, K. J. and Silman, M. R., 2011. The data void in modeling current and future distributions of tropical species. *Global Change Biology*, *17* (1), 626–630.

Ferrier, S., Drielsma, M., Manion, G. and Watson, G., 2002. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales . II . Community-level modelling. *Biodiversity and Conservation*, *11* , 2309–2338.

Ferrier, S. and Drielsma, M., 2010. Synthesis of pattern and process in biodiversity conservation assessment: a flexible whole-landscape modelling framework. *Diversity and Distributions*, *16* (3), 386–402.

Ferrier, S. and Guisan, A., 2006. Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, *43* (3), 393–404.

Ferrier, S., Manion, G., Elith, J. and Richardson, K., 2007. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions*, *13* (3), 252–264.

Ferrier, S., Powell, G. V. N., Richardson, K. S., Manion, G., M, J., Allnutt, T. F., Cameron, S. E., Mantle, K., Burgess, N. D., Faith, D. P., Lamoreux, J. F., Kier, G., Hijmans, R. J., Funk, V. A., Cassis, G. A., Fisher, L., Flemons, P., Lees, D., Lovett, J. C., Rompaey, R. S. A. R. Van, Overton, J. M., Allnutt, F., Fisher, B. L., Lovett, J. O. N. C. and Rompaey, R. S. A. R. V. A. N., 2004. Mapping More of Terrestrial Global Biodiversity for Conservation Assessment. *BioScience*, *54* (12), 1101–1109.

Ferrier, S., 2002. Mapping Spatial Pattern in Biodiversity for Regional Conservation Planning : Where to from Here ? *Systematic Biology*, *51* (2), 331–363.

Field, R., Hawkins, B. a., Cornell, H. V., Currie, D. J., Diniz-Filho, J. A. F., Guégan, J.-F., Kaufman, D. M., Kerr, J. T., Mittelbach, G. G., Oberdorff, T., O'Brien, E. M. and Turner, J. R. G., 2009. Spatial species-richness gradients across scales: a meta-analysis. *Journal of Biogeography*, *36* (1), 132–147.

Fitzpatrick, M. C., Sanders, N. J., Normand, S., Svenning, J.-C., Ferrier, S., Gove, A. D. and Dunn, R. R., 2013. Environmental and historical imprints on beta diversity: insights from variation in rates of species turnover along gradients. *Proceeding of The Royal Society of Biological Sciences*, *280* (1768), 1–8.

Gerstner, K., Dormann, C. F., Václavík, T., Kreft, H. and Seppelt, R., 2014. Accounting for geographical variation in species-area relationships improves the prediction of plant species richness at the global scale. *Journal of Biogeography*, *41* (2), 261–273.

Golicher, D. J., Cayuela, L. and Newton, A. C., 2012. Effects of Climate Change on the Potential Species Richness of Mesoamerican Forests. *Biotropica*, *44* (3), 284–293.

Harte, J., Smith, A. B. and Storch, D., 2009. Biodiversity scales from plots to biomes with a universal species-area curve. *Ecology Letters*, *12* (8), 789–97.

He, K. and Zhang, J., 2009. Testing the correlation between beta diversity and differences in productivity among global ecoregions, biomes, and biogeographical realms. *Ecological Informatics*, *4* (2), 93–98.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. and Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25* (15), 1965–1978.

Jennings, M. D., Hoekstra, J., Higgins, J. and Boucher, T., 2008. A comparative measure of biodiversity based on species composition. *Biodiversity and Conservation*, *17* (4), 833–840.

Keil, P., Schweiger, O., Kühn, I., Kunin, W. E., Kuussaari, M., Settele, J., Henle, K., Brotons, L., Pe'er, G., Lengyel, S., Moustakas, A., Steinicke, H. and Storch, D., 2012. Patterns of beta diversity in Europe: the role of climate, land cover and distance across scales. *Journal of Biogeography*, 1–14.

Kier, G., Mutke, J., Dinerstein, E., Ricketts, T. H., Kuper, W., Kreft, H. and Barthlott, W., 2005. Global patterns of plant diversity and floristic knowledge. *Journal of Biogeography*, *32* (7), 1107–1116.

Koleff, P., Gaston, K. J. and Lennon, J. J., 2003.b Measuring beta diversity for presence-absence data. *Journal of Animal Ecology*, *72* (3), 367–382.

Kraft, N. J. B., Comita, L. S., Chase, J. M., Sanders, N. J., Swenson, N. G., Crist, T. O., Stegen, J. C., Vellend, M., Boyle, B., Anderson, M. J., Cornell, H. V., Davies, K. F., Freestone, a. L., Inouye, B. D., Harrison, S. P. and Myers, J. a., 2011. Disentangling the Drivers of Diversity Along Latitudinal and Elevational Gradients. *Science*, *333* (6050), 1755–1758.

Kreft, H. and Jetz, W., 2010. A framework for delineating biogeographical regions based on species distributions. *Journal of Biogeography*, *37* , 2029–2053.

Leathwick, J. R., Moilanen, a., Ferrier, S. and Julian, K., 2010. Complementarity-based conservation prioritization using a community classification, and its application to riverine ecosystems. *Biological Conservation*, *143* (4), 984–991.

Londoño-Murcia, M. C., Tellez-Valdés, O. and Sánchez-Cordero, V., 2010. Environmental heterogeneity of World Wildlife Fund for Nature ecoregions and implications for conservation in Neotropical biodiversity hotspots. *Environmental Conservation*, *37* (02), 116–127.

Longino, J., Coddington, J. and Colwell, R., 2002. The ant fauna of a tropical rain forest: estimating species richness three different ways. *Ecology*, *83* (3), 689–702.

Manion, G., 2009. A technique for constructing monotonic regression splines to enable non-linear transformation of GIS rasters. *In*: *18th World IMACS/MODSIM Congress*. Cairns, Australia 2507–2513.

McCafferty, W. P., 1998. Ephemeroptera and the Great American Interchange. *Journal of the North American Benthological Society*, *17* (1), 1–20.

McKnight, M. W., White, P. S., McDonald, R. I., Lamoreux, J. F., Sechrest, W., Ridgely, R. S. and Stuart, S. N., 2007. Putting beta-diversity on the map: broad-scale congruence and coincidence in the extremes. *PLoS Biology*, *5* (10), 2424–2432.

Melo, A. S., Rangel, T. F. L. V. B. and Diniz-Filho, J. A. F., 2009. Environmental drivers of beta-diversity patterns in New-World birds and mammals. *Ecography*, *32* (2), 226–236.

Miller, K., Chang, E. and Johnson, N., 2001. *Defining common ground for the Mesoamerican biological corridor*. *Management*. Washington, D.C.: World Resources Institute.

Mokany, K., Harwood, T. D., Overton, J. M., Barker, G. M. and Ferrier, S., 2011. Combining α - and β - diversity models to fill gaps in our knowledge of biodiversity. *Ecology Letters*, *14* , 1043–1051.

Morlon, H., Chuyong, G., Condit, R., Hubbell, S., Kenfack, D., Thomas, D., Valencia, R. and Green, J. L., 2008. A general framework for the distance-decay of similarity in ecological communities. *Ecology Letters*, *11* (9), 904–17.

Mu, Q., Zhao, M. and Running, S. W., 2011. Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment*, *115* (8), 1781–1800.

Nekola, J. and White, P., 1999. The distance decay of similarity in biogeography and ecology. *Journal of Biogeography*, *26* , 867–878.

Ochoa-Ochoa, L. M., Rodríguez, P., Mora, F., Flores-Villela, O. and Whittaker, R. J., 2012. Climate change and amphibian diversity patterns in Mexico. *Biological Conservation*, *150* (1), 94–102.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'amico, J. a., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P. and Kassem, K. R., 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience*, *51* (11), 933–938.

Pausas, J. and Austin, M., 2001. Patterns of plant species richness in relation to different environments: An appraisal. *Journal of Vegetation Science*, *12* (1992), 153–166.

Pearce, J. L. and Boyce, M. S., 2006. Modelling distribution and abundance with presence-only data. *Journal of Applied Ecology*, *43* (3), 405–412.

Rosauer, D. F., Ferrier, S., Williams, K. J., Manion, G., Keogh, J. S. and Laffan, S. W., 2014. Phylogenetic generalised dissimilarity modelling: a new approach to analysing and predicting spatial turnover in the phylogenetic composition of communities. *Ecography*, *37* , 21–32.

Soininen, J., McDonald, R. and Hillebrand, H., 2007. The distance decay of similarity in ecological communities. *Ecography*, *30* (1), 3–12.

Steinitz, O., Heller, J., Tsoar, A., Rotem, D. and Kadmon, R., 2005. Predicting Regional Patterns of Similarity in Species Composition for Conservation Planning. *Conservation Biology*, *19* (6), 1978–1988.

Svenning, J.-C., Fløjgaard, C. and Baselga, A., 2010. Climate, history and neutrality as drivers of mammal beta diversity in Europe: insights from multiscale deconstruction. *The Journal of Animal Ecology*

Swenson, N. G., Anglada-Cordero, P. and Barone, J. a, 2010. Deterministic tropical tree community turnover: evidence from patterns of functional beta diversity along an elevational gradient. *Proceeding of The Royal Society of Biological Sciences*, *278* (1707), 877–84.

Walther, B. and Morand, S., 1998. Comparative performance of species richness estimation methods. *Parasitology*, *116* , 395–405.

Wiersma, Y. F. and Urban, D. L., 2005. Beta Diversity and Nature Reserve System Design in the Yukon, Canada. *Conservation Biology*, *19* (4), 1262–1272.

Williams, K. J., Belbin, L., Austin, M. P., Stein, J. L. and Ferrier, S., 2012. Which environmental variables should I use in my biodiversity model? *International Journal of Geographical Information Science*, 1–39.

Williams, K. J., Ferrier, S., Rosauer, D., Yeates, D., Manion, G., Harwood, T., Stein, J., Faith, D. P., Laity, T. and Whalen, A., 2010. *Harnessing Continent-Wide Biodiversity Datasets for Prioritising National Conservation Investment*. *Communities*. Canberra, Australia.

Wu, F., Yang, X. J. and Yang, J. X., 2010. Additive diversity partitioning as a guide to regional montane reserve design in Asia: an example from Yunnan Province, China. *Diversity and Distributions*, *16* (6), 1022–1033.

Yesson, C., Brewer, P. W., Sutton, T., Caithness, N., Pahwa, J. S., Burgess, M., Gray, W. A., White, R. J., Jones, A. C., Bisby, F. a and Culham, A., 2007. How global is the global biodiversity information facility? *PloS One*, *2* (11), e1124.

3 CHAPTER 3

Incorporating beta diversity into estimates of tree diversity loss in the Mesoamerican region

3.1 INTRODUCTION

The past decade can be characterized by increased recognition of the importance of biodiversity to human existence and the need to maintain a natural, healthy environment (Margules & Pressey 2000). Population growth and global financial crises exacerbate the demands on natural resources. This drives the need to prioritize areas for efficient allocation of conservation efforts. The development of methods to select priority areas is an expanding field of research that has given rise to various approaches and metrics to compare the biological importance of different areas (Funk & Fa 2010). It is generally agreed that conservation areas should be designed to ensure long-term persistence of the ecological elements and processes that maintain biodiversity (Fairbanks et al. 2001) and give rise to provision of ecosystem services (Pressey et al. 2007; Funk & Fa 2010). It is assumed that increasing the number of reserves would capture more diversity. However, this begs the question of whether additional reserves would be capturing different suites of species or if they would just encompass more of the same species. The establishment of representative conservation areas, therefore, depends on the degree of species compositional heterogeneity, or beta diversity, within the region of interest (Wu et al. 2010).

Beta diversity has been interpreted in many different ways since the term was first defined, and some authors have attempted to concretely define a universal concept (Tuomisto & Ruokolainen 2006; Tuomisto 2010a; Tuomisto 2010b) or to simplify the many ideas into a workable framework (Anderson et al. 2011). In his seminal papers (Whittaker 1960, 1972), Whittaker defines beta diversity in two ways: (i) as the ratio between regional (gamma) diversity to local (alpha) diversity, and (ii) the variation in species composition along habitat gradients, usually measured through pair-wise comparison of focal areas. These definitions followed from the concept of niche theory which argues that species occupy different positions along habitat gradients based on differences in their utilization of resources (Whittaker 1972). In this paper we focus on the second of Whittaker's definitions. We use the concept that beta diversity or spatial compositional turnover, is the variation in species composition among localities (Whittaker 1972; Ferrier 2002; Koleff & Gaston 2002).

In order to measure beta diversity, many indices have been developed. Koleff et al. (2003) provides an extensive review of many of the beta diversity indices that have been applied to the study of beta diversity patterns. That review provides a clear conceptualization of beta diversity indices by re-expressing them according to three components, where component 'a' equals the number of shared species between the two samples being compared, 'b' equals the number of species found only in the focal sample (species loss) and 'c' equals the number of species found only in the neighbouring sample (species gain). This

conceptualization provides a simple framework for comparing the many indices and deciding on an approach for assessments comparing different areas and aimed at prioritizing conservation efforts.

Most conservation prioritization efforts at regional or global levels focus on measuring endemism of individual taxa or groups of species (Muller et al. 2003; Brooks et al. 2006; Morawetz & Raedig 2007; Funk & Fa 2010). These studies assume that the focal taxa are appropriate surrogates for more general biodiversity patterns, but studies have found weak support for this assumption (Leathwick et al. 2010). Other studies focus on prioritizing areas based on inferred or modelled species richness (Anderson et al. 2008; Golicher et al. 2011). These studies represent efforts to shift the focus from modelling individual entities to modelling collective properties of biodiversity, such as species richness (Ferrier 2002; Ferrier et al. 2002). Such efforts include the many applications of individual species distribution models to predict the species richness and species composition of a given area (Ferrier & Guisan 2006; Benito et al. 2013). Despite the utility of species distribution models in identifying patterns of species richness across large scales, species richness itself may be inadequate for use in regional conservation planning since it does not provide information on the level of complementarity (Margules & Pressey 2000) or representativeness of a set of areas (Ferrier et al. 2004). Whereas species richness accounts only for the number of species in comparing areas, beta diversity or compositional turnover accounts for the identities of the species in the areas being compared (Jennings et al. 2008). By considering species identities, beta diversity introduces a way to determine the degree to which an area is similar or distinct from other areas (Jennings et al. 2008). In this way beta diversity provides an important measure for comparisons between areas. Importantly, this measure can be incorporated into the planning process of conservation initiatives that seek to maximize biodiversity benefits.

Land use change caused by human activities is one of the major threats to biodiversity through destruction of natural habitat (Jetz et al. 2007). Tropical regions harbour the majority of the Earth's species and face rapid land use change (Feeley & Silman 2011). For example, in Madagascar, 9.1% of certain invertebrate and plant species were estimated to have been lost between 1950 and 2000 as a result of deforestation (Allnutt et al. 2008). In tropical and subtropical regions, land use change will potentially cause severe range losses in taxa such as birds (Jetz et al. 2007). In these rapidly changing tropical landscapes, land use change may affect species' ability to survive climate change as patterns of land use may prevent species from migrating to climatically suitable habitat (Feeley & Silman 2010). It is therefore important to understand the effect of land use change on species diversity. Any such analyses of anthropogenic effects on biodiversity must include the magnitude and spatial patterning of land use changes.

In this chapter, I integrate the GDM approach described in **Chapter 2** with the Global Biodiversity Model (GLOBIO3), which was developed to assess anthropogenic impacts on biodiversity, such as those arising from land use change (Alkemade et al. 2009). Using the GDM model from **Chapter 2**, I extend the method of Allnutt et al. 2008 (based on Ferrier et al. 2004) by introducing a continuous measure of habitat condition called Mean Species Abundance (MSA) used in the GLOBIO3 model, to estimate the proportion of habitat

and tree species loss in the region up to the year 2000. We use this assessment to indicate which ecoregions have been most affected by land use change and may therefore warrant prioritization in conservation efforts. This approach of combining predicted species turnover derived from the GDM model and the MSA measure of habitat condition represents a novel application of the SAR to estimating species loss.

The objective of this chapter is to apply the GLOBIO framework and GDM approach to assess the impact of land use change on tree species diversity within the region, to answer the following questions:

1. What proportion of habitat and species has been lost as a result of land use changes up to the year 2000?
2. What is the effect of land use change on both areas that are compositionally more similar to other areas in the region (low beta diversity), and those that are less similar (high beta diversity)?
3. How does habitat and species loss vary among ecoregions?

3.2 METHODS

3.2.1 Study area

Mainland Mesoamerican encompasses the five southern states of Mexico (Campeche, Chiapas, Quintana Roo, Tabasco, and Yucatan) and all Central American countries (Belize, Guatemala, El Salvador, Honduras, Nicaragua, Costa Rica and Panama). The region encompasses a land area of 768,990 km^2 and is home to a population of about 45 million (half of which is thought to live below the poverty line), which continues to grow at approximately 2% per year (Miller et al. 2001). Developmental pressures arising from poverty and an increasing population are drivers of the environmental degradation occurring over recent years (Conservation International 2004). The region is threatened by intense land use change (Golicher et al. 2012) including a deforestation rate of about 2.1% per year since measured in 1999 (Miller et al. 2001).

The region has been classified into 35 ecoregions including lowland and mountain forests, grasslands, pine savannahs, and mangroves. These were delineated based on known mapping of floristic and faunal distributions, vegetation types, and data and consultation from experts (Olson et al. 2001). These ecoregion units have been used as a basis for conservation planning efforts within Mesoamerica, such as in gap analyses to assess how much they are represented within existing protected area networks (Londoño-Murcia et al. 2010; Xiao-jun et al. 1999). I use these currently-defined ecoregions as the basic units to conduct the analysis of species loss.

3.2.2 Data

3.2.2.1 Species Data

The dataset used for this analysis consists of over 435,000 geo-referenced specimen records from the GBIF, representing 2,561 species of mostly woody trees and shrubs (height > 4m; most 5-10cm dbh). The dataset has undergone taxonomic standardization using the R package TAXONSTAND Version 1.2, which standardizes plant names using The Plant List (http://www.theplantlist.org) to retrieve information about each species taxonomic status (Cayuela et al. 2012). In addition, the species data set was filtered to remove obvious duplicate records, records which had unknowns at any of the taxonomic levels (family, genus, species), records with coordinates rounded to the nearest degree, records with uncertainty values of > 10,000 meters, and records of any grasses or sedges. All processing was done in R v. 2.13.2 (R Development Core Team 2011).

3.2.2.2 Climatic data:

Bioclimatic predictor grids were derived from the WorldClim bioclimatic variables dataset (Hijmans et al. 2005); www.worldclim.org). Bioclimatic variable grids were downloaded from Worldclim at 30 arc-seconds (~1km) resolution. Since the grid cell size of our study was defined at 0.1° (~10km), the above bioclimatic grids had to be resampled to that resolution. A new grid for each bioclimatic variable was derived by averaging the values of all pixels within the 30 arc-second bioclimatic grids to the desired 10km resolution. An evapotranspiration grid was also derived at 0.1° (~10km) resolution from the MODIS global terrestrial evapotranspiration data set (MOD16) (Mu et al. 2011).

3.2.3 Analytical framework

In this study, Generalized Dissimilarity Modelling (GDM) is combined with a measure of habitat condition, known as Mean Species Abundance (MSA), in order to estimate species loss after land use change. This GDM-MSA approach applies a general methodological framework that links spatial species turnover (beta diversity) with species loss through land use change. This approach involves three major steps (**Figure 3.1**):

(i) derive an appropriate habitat condition map
(ii) estimate Effective Habitat Area (EHA) after land use change up to 2000
(iii) apply the species-area relationship (SAR) to estimate the proportion of species retained

These steps are explained in more detail in the following sections.

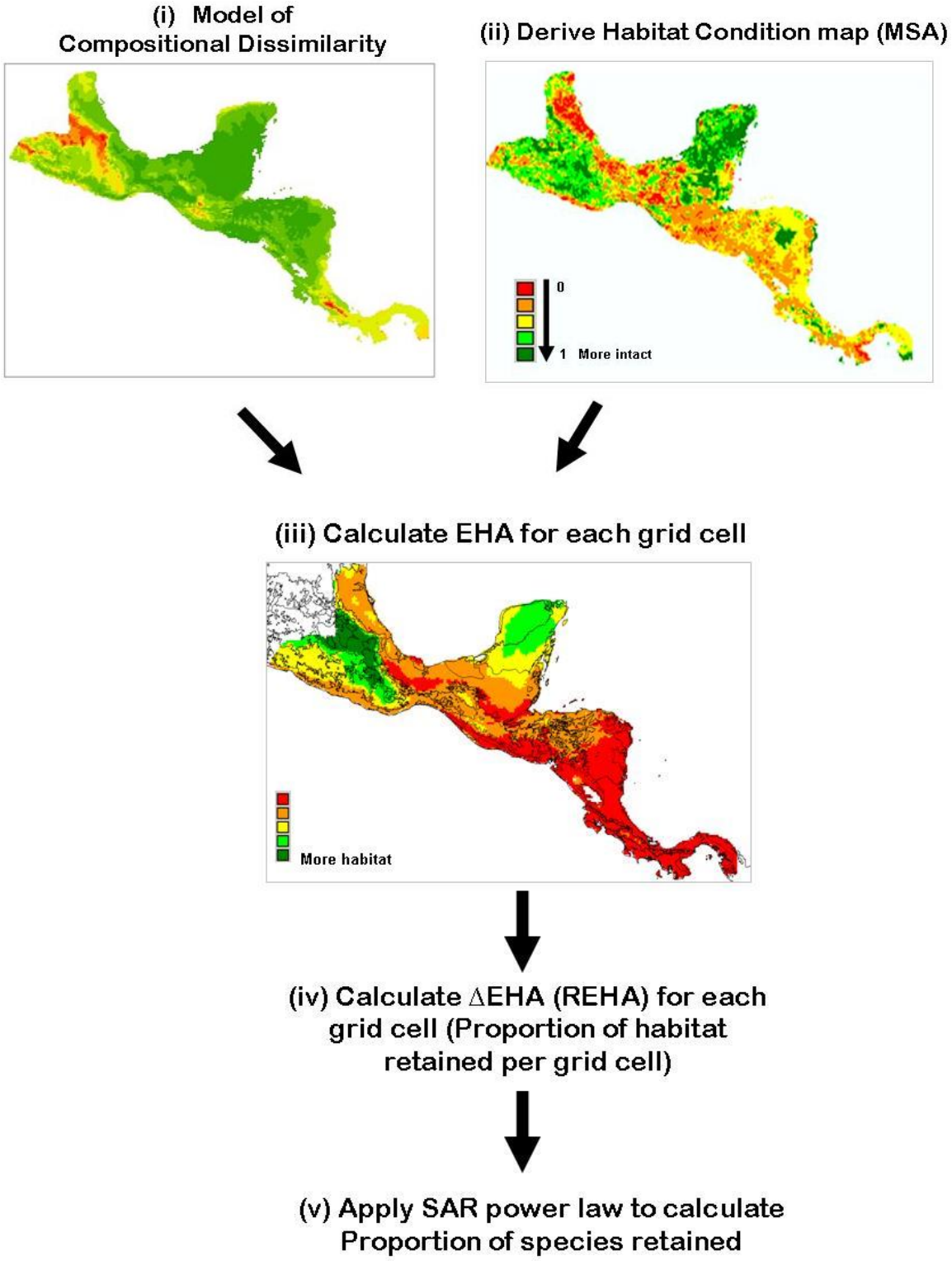


Figure 3.1 Diagram showing major analysis steps involved in the GDM approach for estimating proportion of species retained or lost up to a given baseline year.

3.2.3.1 Deriving an appropriate Habitat Condition map

In previous studies, land use change has been incorporated either by the use of a compound index such as the human footprint index (Broennimann et al. 2006) or by the binary classification of areas as forested or non-forested (Allnutt et al. 2008). In this analysis, we introduce a continuous measure of habitat condition that represents the magnitude of the effect on biodiversity after an area has experienced land use change. The Mean Species Abundance (MSA) measure is the main measure of biodiversity used by the GLOBIO3 model that was developed to assess the impact of anthropogenic and climatic drivers on biodiversity (Alkemade et al. 2009). GLOBIO3 describes biodiversity as the remaining mean species abundance (MSA) of original species in a disturbed habitat, relative to their abundance in pristine or undisturbed vegetation (Alkemade et al. 2009). The pristine habitat is assumed be undisturbed by human activities for a prolonged period (Alkemade et al. 2009). The value of MSA for different land use types varies from 0.1 for intensive agriculture to 1 for primary forest (**Figure 3.2**), and is measured relative to the assumed undisturbed condition.

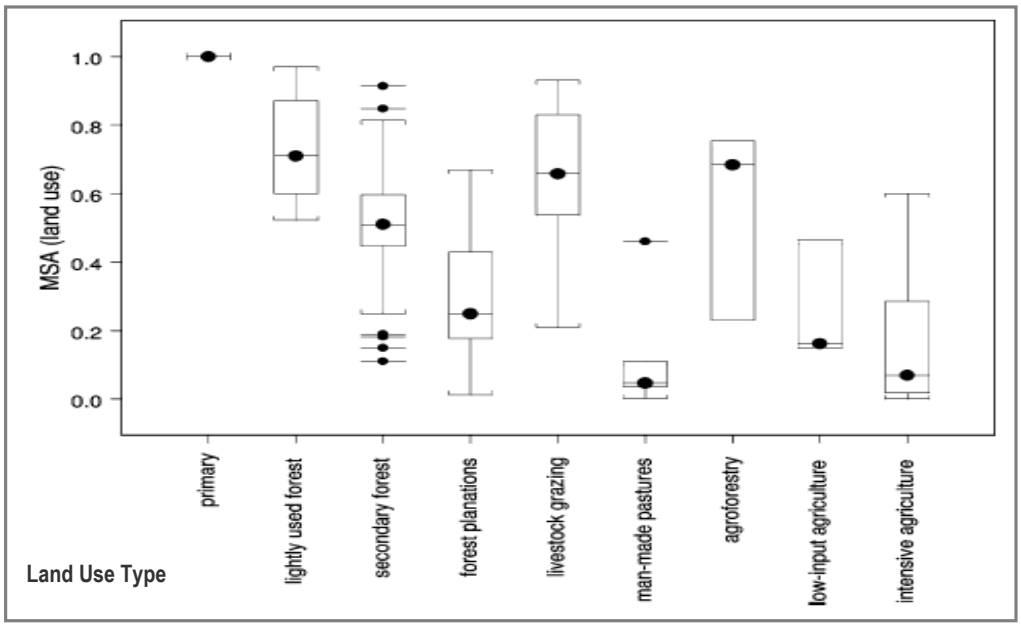


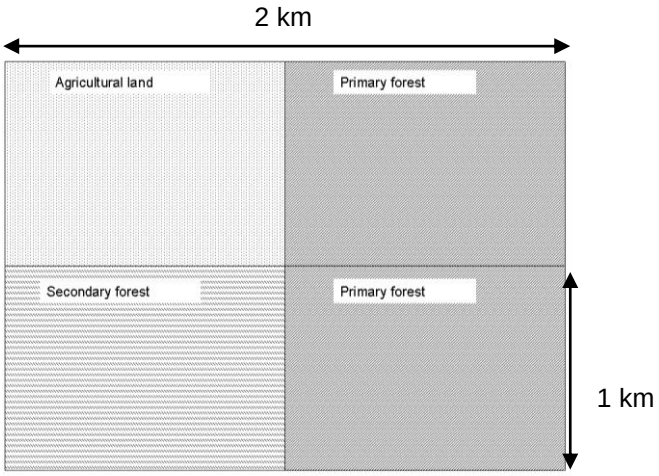
Figure 3.2 The MSA values for each land use type. Adapted from (Alkemade et al. 2009).

The MSA measure was developed based on a meta analysis of studies documenting the effect of disturbed habitat on various taxa, including vertebrates and plants, in contrast to undisturbed habitat (Alkemade et al. 2009). It can be used to assess the effect of different impacts, such as land use change and climate change, on plants and vertebrates. In this case, we assessed only the effect of land use change on plants.

For each 0.1 degree grid cell in our study, the land use types contained in the cell were identified. The land use types identified in each cell are based on the GlobCover 2009 land cover product (Alkemade et al. 2009). The land area occupied by each land-use type within the cell was calculated. Each land use type has an MSA

value based on the effect on biodiversity associated with that land use type as evidenced by the meta-analysis studies (Alkemade et al. 2009). The product of that MSA value and the land area covered by a land use type equalled the fractional MSA value for that land use type within the cell. The total MSA value for the entire grid cell was calculated as the sum of fractional MSA values across all land use types found in the grid cell, weighted by the cell area (see **Figure 3.3** for an example).

In this way, an MSA value for each grid cell in our study area was obtained. This serves as the input raster of habitat condition for the next step: estimating proportion of original species retained.



Land use type	**MSA**
Agricultural land	0.1
Primary Forest	1
Secondary Forest	0.5

MSA for entire cell = [(area of agricultural land x agriculture MSA) + (area of primary forest x primary forest MSA) + (area of secondary forest x secondary forest MSA)] / Total cell area

= [(1x0.1) + (2x1) + (1x0.5)] / 4
= (0.1 + 2 + 0.5) / 4
= 0.65

Figure 3.3 An example of calculating MSA value for a 4km^2 grid cell. The sum of fractional MSA values across the three land use types in this cell weighted by the total cell area gives the final MSA value for the entire cell. MSA values assigned to different land use types according to Figure 3.2 (Alkemade et al. 2009).

3.2.3.2 Estimate effective habitat area (EHA) after land use change

In **Chapter 2**, EHA was calculated for the assumed pristine condition. In this chapter, EHA is calculated after land use change, that is, after the habitat has been disturbed following the same equation (1) in **Chapter 2**. Allnutt et al. (2008) quantified the habitat condition (C_j) of each cell in a binary fashion, as either forested or deforested (0 or 1, respectively). In contrast to this, we introduce a continuous measure of habitat condition by using MSA as the measure of habitat condition. EHA for each cell in this scenario is defined as the sum of its similarities ($\sum S_{ij}$) multiplied by its MSA value.

3.2.3.3 Applying the SAR power-law function to estimate proportion of species retained

Following the methodology described in Ferrier et al. (2004) and Allnutt et al. (2008), the species-area relationship (SAR) (Arrhenius 1921) can be combined with predicted spatial patterns of dissimilarity to predict the proportion of biodiversity retained under a certain land use scenario. This method modifies the SAR power-law function to estimate the total proportion of species retained after land use change relative to the original condition.

The SAR power-law function describes the relationship between number of species (S) and area (A) as:

$$S = cA^z \qquad \text{(Equation 2)}$$

where c and z are constants (Allnutt et al. 2008). This function can be modified to predict the proportion of species that will be retained in a region over time if the region's habitat is reduced to a proportion of the original area (Allnutt et al. 2008). The relationship can be re-expressed as:

$$S_{retained} / S_{original} = (A_{retained} / A_{original})^z. \qquad \text{(Equation 3)}$$

Using equation (2), we can estimate the proportion of species retained ($S_{retained}$) once we know the original habitat area ($A_{original}$) and the new habitat area after land use change ($A_{retained}$). Ferrier et al. (2004) and Allnutt et al. (2008) extend this concept to incorporate beta diversity by using GDM to estimate the proportion of species that are shared between a pair of study sites, in this case, between pairs of grid cells (Allnutt et al. 2008).

Allnutt et al. (2008) modifies the SAR power-law function thus:

$$\text{Total proportion species retained } (P) = \frac{\sum_{i=1}^{n}\left(\left(\sum_{j=1}^{n} s_{ij}c_j \Big/ \sum_{j=1}^{n} s_{ij}\right)^{z}\left(1 \Big/ \sum_{j=1}^{n} s_{ij}\right)\right)}{\sum_{i=1}^{n}\left(1 \Big/ \sum_{j=1}^{n} s_{ij}\right)} \qquad \text{(Equation 4)}$$

(A = Term A, B = Term B)

where Term A is change in Effective Habitat Area (EHA): the proportion of habitat remaining for species originally occurring in cell *I;*

Term B adjusts for compositional overlap between cells

s_{ij} = predicted compositional similarity between cells i and j from GDM model

c_j = habitat condition of cell j (MSA value)

z = species-area exponent (assumed constant at 0.25)

The ratio of the undisturbed EHA to the MSA-scenario EHA is the final change in EHA after land use change. Term 'A' (change in EHA) as described in equation (4) above is the quantity to which the exponent is applied to estimate the proportion of species retained after land use change. It is an estimate of the proportion of habitat remaining for species originally occurring in a given focal cell *i* and is equivalent to $A_{retained}$ / $A_{original}$ (Ferrier et al. 2004). In this thesis, I refer to the change in EHA, as REHA (remaining habitat for original species).

By applying the exponent (z=0.25), the proportion of habitat remaining for the species originally occurring in each cell is used to estimate the proportion of *species* remaining in each cell (Ferrier et al. 2004). The rest of the formula combines individual cell estimates into an overall estimate of the proportion of species retained within the region of interest (Ferrier et al. 2004). The weights applied in the formula, via Term B, adjust for compositional overlap between cells. They ensure that more distinctive cells (i.e., cells that are similar to few other cells) contribute as equally as cells that are similar to many other cells, in deriving a final *P* value for the region (Allnutt et al. 2008).

The output of the GDM modelling process gives predicted dissimilarity values (d_{ij}). For simplicity, in equation (4), the similarity equivalent is used, so that predicted similarity between cells *i* and *j* is calculated as $s_{ij} = 1 - d_{ij}$ (Allnutt et al. 2008). All calculations were carried out using the bespoke software *Muru: Pathways to Resilience* developed by Tom Harwood at CSIRO (Commonwealth Scientific and Industrial Research Organization).

Allnutt et al. (2008) quantified species loss across one scale, the island of Madagascar. In this analysis, we quantify species loss across the entire Mesoamerican region but also subdivide it into ecoregions and quantify species loss at this scale, in order to compare the effect of land use change on these conservation units. The predicted species loss given by this analysis refers to biodiversity "committed to extinction", since the loss may take time to be realized (Allnutt et al. 2008). This analysis uses MSA values derived from Global Land Cover data for the year 2000. Therefore, this analysis can be said to project the proportion of tree species that were "committed to extinction" given the land use changes that had taken place up to that year.

3.3 RESULTS

3.3.1 Spatial pattern of habitat condition

This analysis introduced MSA as a continuous measure of habitat condition or intactness for each grid cell. MSA values varied from 0 (highly disturbed habitat) to 1 (slightly disturbed habitat). Overall, more intact

habitat occurred to the west of the study area and in the Yucatan peninsula, in Mexico. The southern areas of Mexico and the rest of Central America showed less intact habitat (**Figure 3.4**).

At the regional level, across all grid cells, there was a significant but weak correlation between a grid cell's protected status and its habitat condition ($r = 0.38$, $n = 9{,}543$, $p < 0.005$) (**Figure 3.5**). However, this belies the variation between countries with respect to this correlation (**Figure 3.6** & **Table 3.2**). At the national level, Guatemala, Belize and Panama showed that the habitat condition of sites (grid cells) within these countries was highly correlated with their protected status. Mexico and El Salvador showed the lowest correlation values indicating that the habitat condition of sites within these countries did not depend on protected status. Surprisingly, Costa Rica showed the third lowest correlation value. One reason for this could be that protected areas were expanded to include or established in highly disturbed areas after 2000. Consequently, the mean MSA within protected areas in Costa Rica would decrease. In Mexico, however, the low correlation seemed more likely to result from areas of high habitat condition (high MSA) still remaining outside of protected areas as of 2000 (**Figure 3.4**).

In general, Mexico, Belize and Costa Rica showed the highest overall mean MSA values across all countries, while El Salvador had the lowest (**Table 3.1**). Nicaragua and Honduras showed the same overall mean MSA values, as did Panama and Guatemala. Based only on overall mean MSA, it would seem that Mexico, Belize, Costa Rica, Nicaragua and Honduras had more intact habitat remaining. However, when the MSA values within and outside protected areas for each country were examined, some differences emerged. In this case, Honduras showed more intact habitat outside protected areas than Costa Rica, while Nicaragua showed the third lowest habitat condition outside protected areas (**Table 3.2**). Costa Rica shows low habitat condition within protected areas; however, this could reflect the expansion or establishment of new protected in highly disturbed areas.

When mean MSA values were calculated for each ecoregion, a clear pattern emerged showing that most ecoregions along the central Pacific coast had below-average MSA values (**Figure 3.7**). The mean MSA value across all ecoregions was 0.522. Most Dry, Montane and Moist Forests ecoregions had below average MSA values. Ecoregions with some of the lowest MSA values representing those that were highly disturbed included Chiapas Depression Dry Forests, Panamanian Dry Forests, Veracruz Moist Forests, Central American Dry Forests, Isthmian-Pacific Moist Forests, and Central American Pine-Oak Forests (**Figure 3.7**). The ecoregions with highest habitat condition values were those in Mexico to the west of the study area and in the Yucatan peninsula. The Talamancan Montane Forests and Eastern Panamaian Montane Forests ecoregions also had relatively high habitat condition vaues. This is because most of these ecoregions are under protection. Most of the remaining ecoregions showed intermediate levels of disturbance.

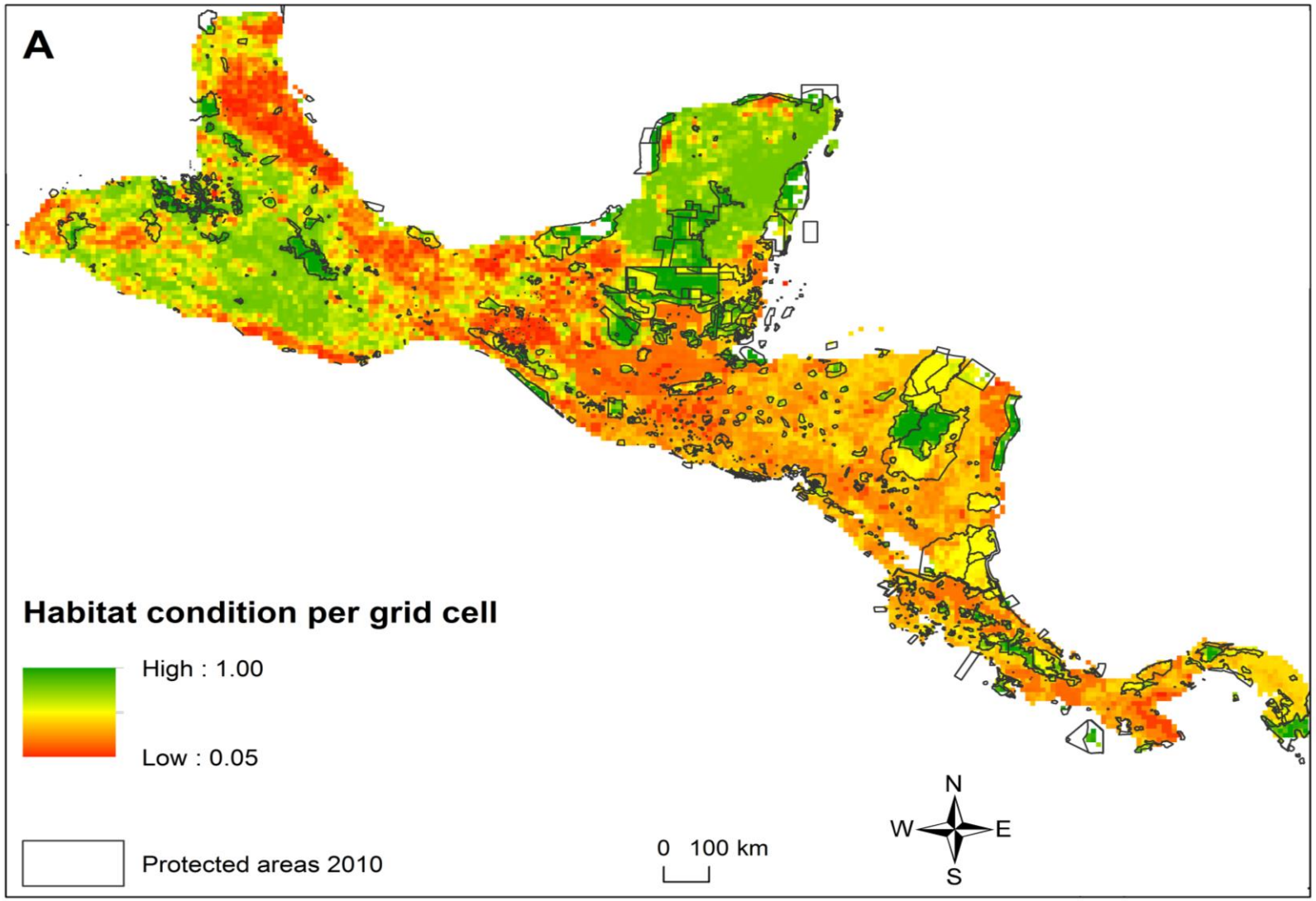


Figure 3.4. Map of habitat condition for each 0.1 degree grid cell as measured by MSA (Mean Species Abundance). High values represent slightly disturbed habitat; low values represent highly disturbed habitat. Polygons represent protected area boundaries. Land cover data from Global Land Cover 2000.

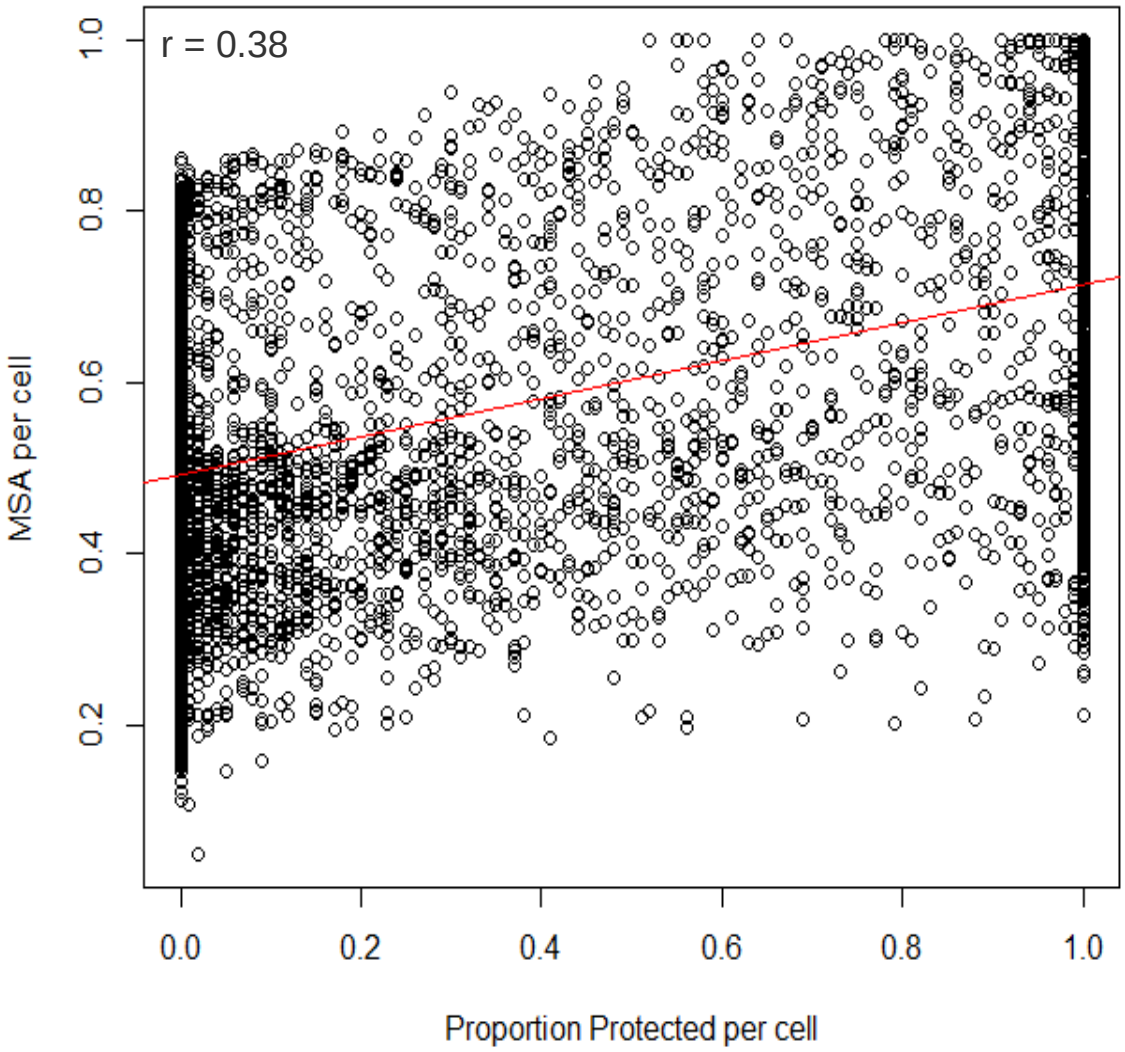


Figure 3.5. Scatterplot showing a significant but weak correlation between the study sites' protected status and their habitat condition (MSA) values.

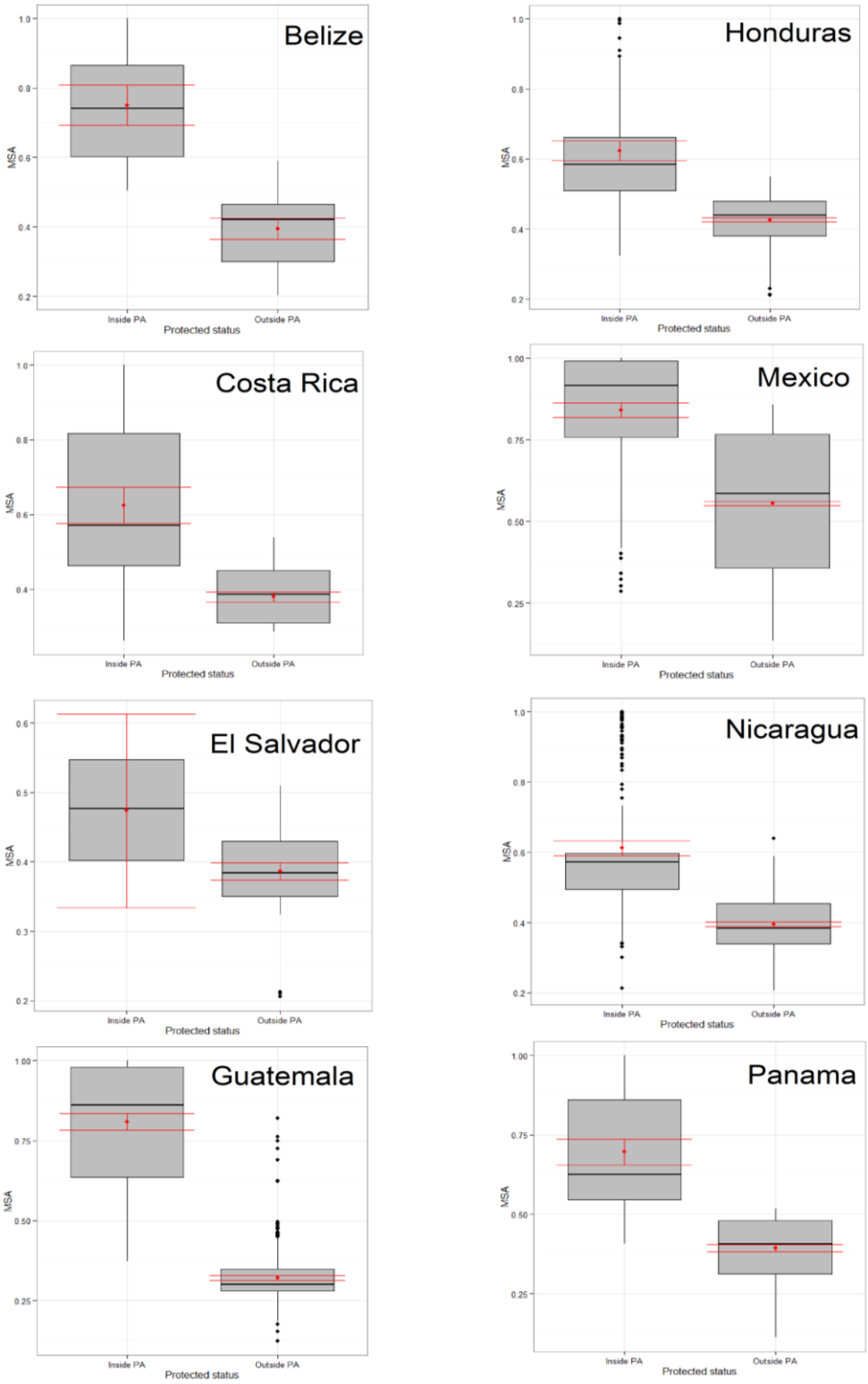


Figure 3.6. Boxplot illustrating variation in habitat condition, as measured by MSA, inside and outside protected areas for each country.

Table 3.1. Table showing mean habitat condition values for each country in the study region (sorted in descending order).

Country	Mean MSA
MEXICO	0.59
BELIZE	0.56
COSTA RICA	0.49
NICARAGUA	0.48
HONDURAS	0.48
PANAMA	0.47
GUATEMALA	0.47
EL SALVADOR	0.40

Table 3.2. Table showing correlation between protected status and habitat condition for all grid cells in each country (sorted in descending order by correlation value), as well as the mean habitat condition values within and outside protected areas for each country.

Country	Pearson's r	Mean MSA in PA	Mean MSA outside PA
Guatemala	0.83	0.781	0.349
Belize	0.79	0.735	0.461
Panama	0.68	0.643	0.408
Honduras	0.60	0.610	0.430
Nicaragua	0.59	0.604	0.403
Costa Rica	0.57	0.598	0.420
El Salvador	0.38	0.472	0.391
Mexico	0.36	0.796	0.564

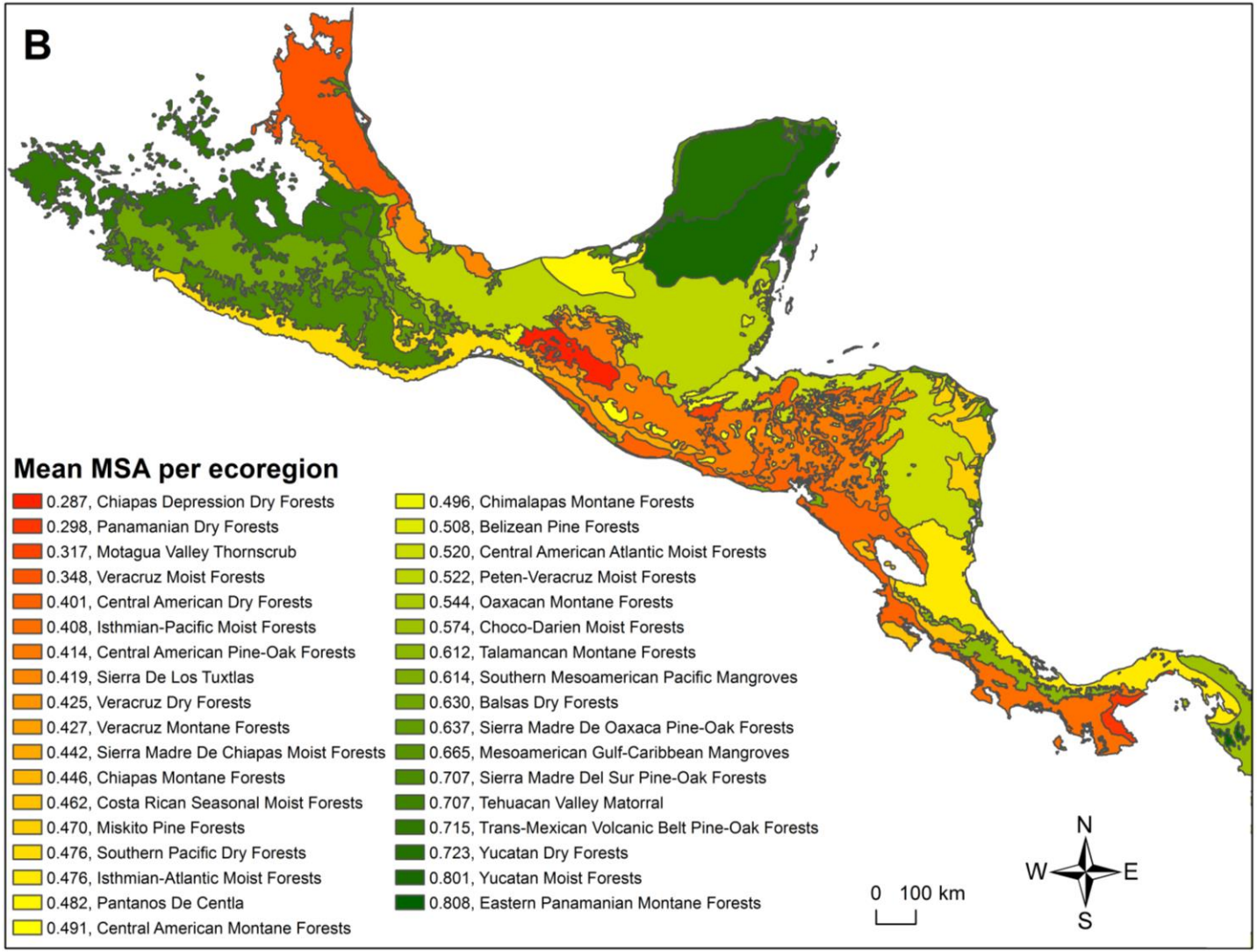


Figure 3.7. Map of mean MSA (mean species abundance) for each ecoregion. High values represent slightly disturbed habitat and low values represent highly disturbed habitat.

3.3.2 Effective Habitat Area (EHA) after land use change

The outputs of the EHA calculation after land use change (see **Methods, Section 3.2.3.2**) are shown for each grid cell in **Figure 3.8**, and for each ecoregion in **Figure 3.9**. Since EHA is calculated as the sum of the product of the similarity and the condition of a cell (Sij x Cj), the EHA for each grid cell is affected by the decrease in habitat condition values (MSA values) as a result of habitat loss or degradation. The predicted similarity remains the same; it is only the habitat condition value that changes. After land use change, habitat condition values can only decrease to values less than one, with a value of one indicating the presumed pristine condition. Hence, the overall EHA values also decrease as a result of land use change. A cell or ecoregion with a high value represents an area that is dissimilar to a larger proportion of the total habitat than a cell or ecoregion with a lower value. These areas, therefore, might indicate areas containing species assemblages that may not occur elsewhere in the region. However, it is important to note that cells or ecoregions similar in value do not necessarily contain the same set of species.

The spatial pattern of EHA after land use change was similar to the pattern at the pristine condition (see **Chapter 2, Figure 2.2**). However, the proportion of EHA for each grid cell almost doubled after land use change (**Figure 3.8**). At the ecoregion scale, the pattern of ecoregions showing low EHA values remained the same as at the pristine condition. In both cases the values increased as a result of habitat loss. It was evident from comparing **Figure 3.7** and **Figure 3.9**, that the most extensive ecoregions with high MSA (low habitat loss) and low EHA were the Dry forests and Pine-Oak forests of Mexico. In contrast, ecoregions with low EHA and low MSA (high habitat loss) were mostly those located in Costa Rica and Panama to the south of the study region, but also included Central American Montane Forests, Veracruz Moist Forests, Veracruz Montane and Veracruz Dry Forests. The ecoregions with high EHA and low MSA were mostly located in the central part of the study region on both the Atlantic and Pacific coasts. Finally the ecoregions with high EHA and high MSA were confined to the Yucatan peninsula.

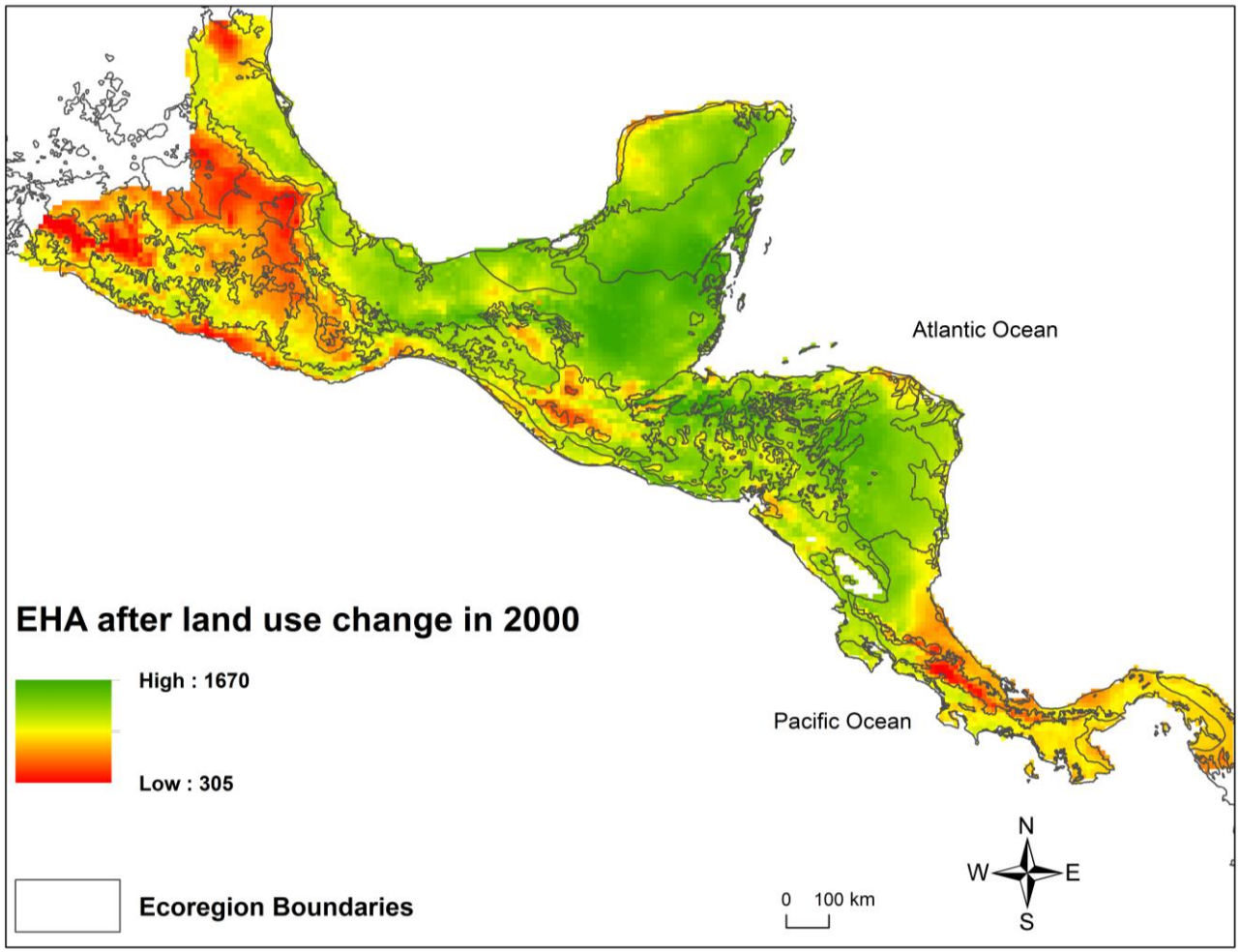


Figure 3.8. EHA for each grid cell after land use change up to 2000.

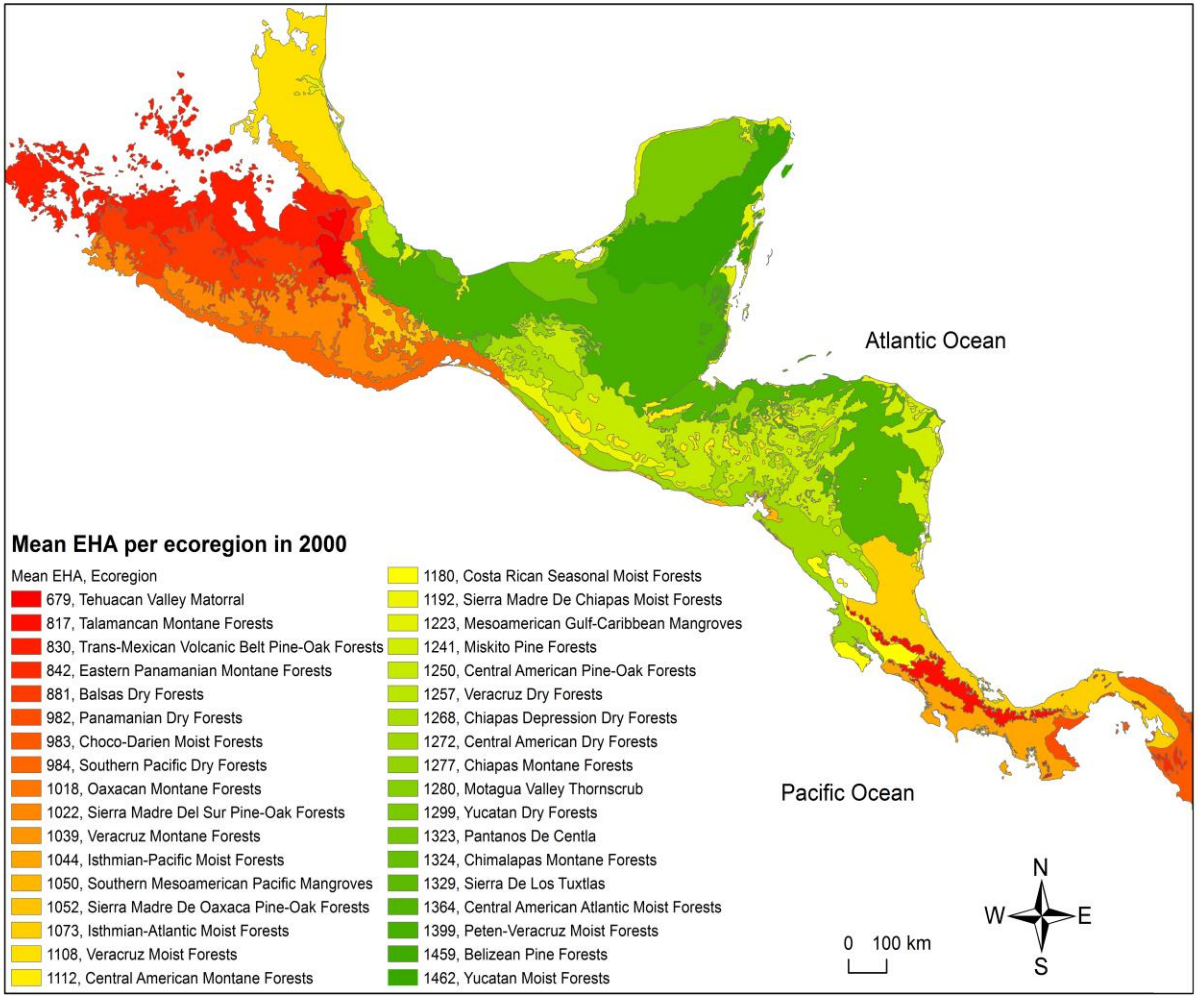


Figure 3.9. Mean EHA per ecoregion, after habitat loss is accounted for.

3.3.3 Land use changes up to year 2000 reduced available habitat by up to 50%

In this analysis, the land use data were used as an indicator of habitat condition and incorporated with the predicted dissimilarity derived from the GDM model (see **Methods, Section 3.2.3.2**) to produce an estimate of the proportion of Remaining Effective Habitat Area (REHA) after land use change for each grid cell in the study region.

This value represents the proportion of habitat remaining for the original species occurring in each cell, at the year 2000. Mean habitat loss across the region was 46.7%. REHA is lowest at around 50% in the southern ecoregions (**Figure 3.10**) because these areas either have low MSA or have low EHA. This property of low MSA or high dissimilar EHA means that the habitat remaining for original species is much lower in the southern ecoregions. The Talamancan Montane Forests, Choco-Darien Moist Forests and Eastern Panamanian Montane Forests, although characterised by relatively high MSA values, also displayed above average dissimilar EHA values. This means that although they have experienced less habitat loss, the effect is great because they harbour highly dissimilar habitat.

Variation in REHA across the region is small, with less than 10% difference between the ecoregions with the highest proportion of REHA to those with the lowest proportion (**Figure 3.11**). This means the effect of habitat loss has been similar across all ecoregions, even in those that have high EHA (low beta diversity). This is illustrated by comparison of two ecoregions: Peten-Veracruz Moist Forest and Central American Atlantic Moist Forests (see **Figure 3.11)**. These ecoregions have high EHA (low beta diversity) but high land use impact as measured by MSA (**Figure 3.7**). This could indicate that the high land use impact alone is driving the lower EHA remaining in these areas. However, in this study, REHA is the proportion of habitat remaining for *original* species. This means that the available habitat for original species occurring in these areas has decreased. Therefore, habitat loss is having a great effect on areas of high EHA (low beta diversity) by driving the loss of originally occurring species just as it is in areas of low EHA or high beta diversity. This means that areas of high EHA were only marginally less affected by habitat loss than areas of low EHA.

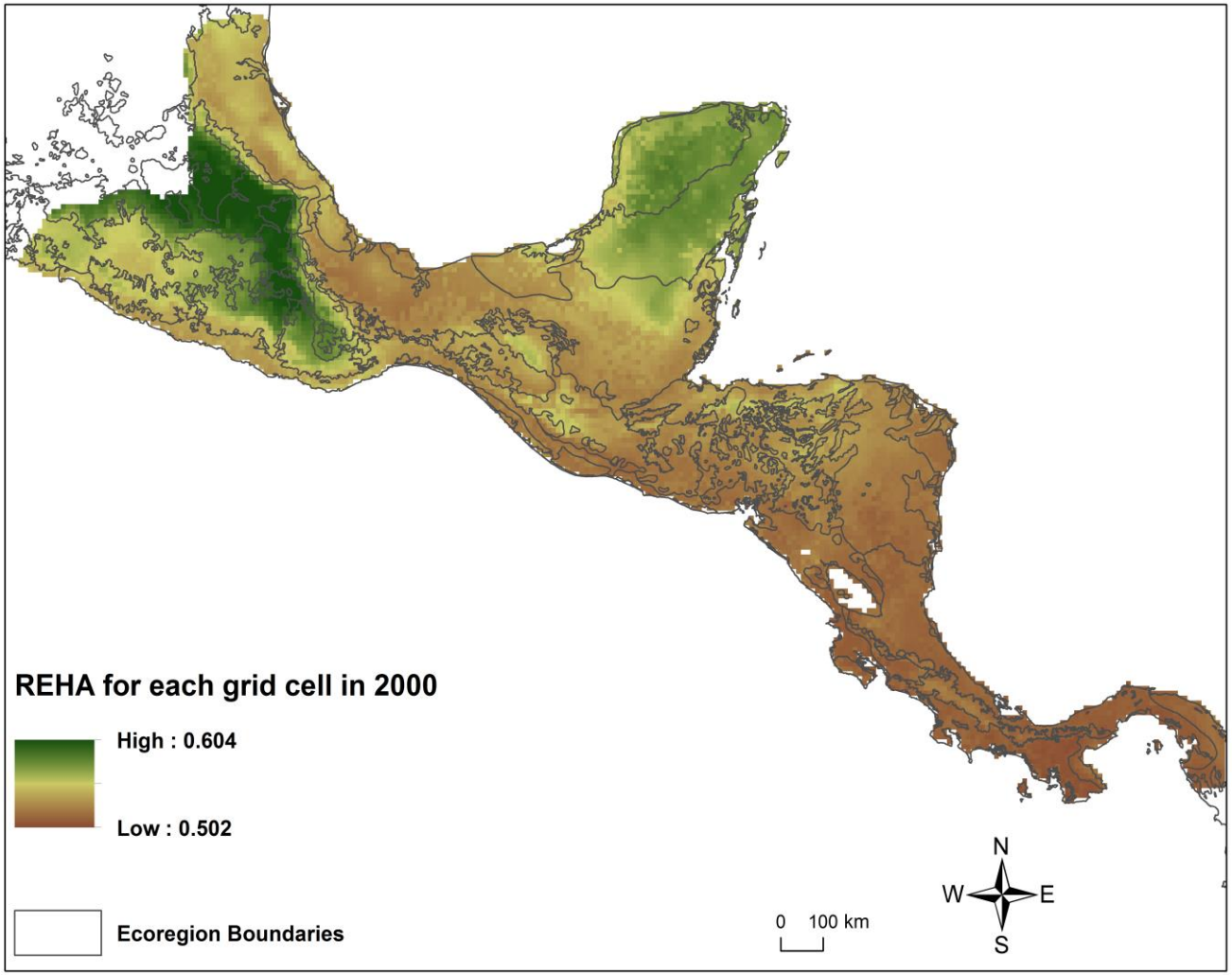


Figure 3.10. Map of Remaining Effective Habitat Area (REHA) for each grid cell. REHA represents the proportion of habitat remaining for originally occurring species.

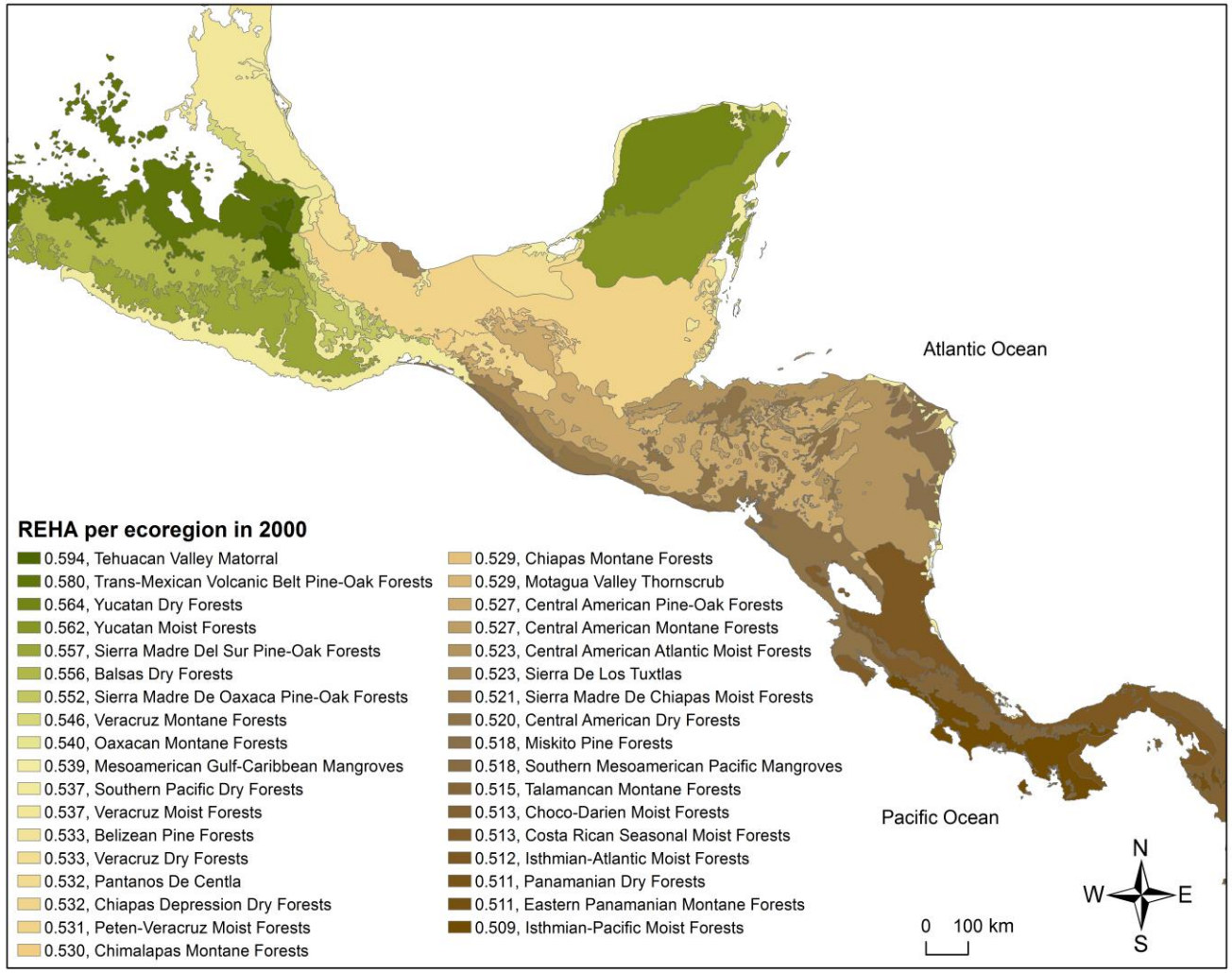


Figure 3.11. Map of mean Remaining Effective Habitat Area (REHA) for each ecoregion. REHA represents the proportion of habitat remaining for originally occurring species.

3.3.4 High land use impact translates to 15% species loss in southern Mesoamerica

Across the entire region of Mesoamerica, 85.7% of tree species were estimated to have been retained, indicating that an estimated 14.3% of species were predicted to have been lost owing to land use change. The estimated proportion of species retained per ecoregion decreased from north to south, across the study region (**Figure 3.12**).

The southernmost ecoregions had the highest proportion of estimated species loss. The Isthmian-Pacific Moist Forests and Eastern Panamanian Montane Forests had incurred the most loss with an estimated 15.5% species loss, followed by Isthmian Atlantic Moist Forests and Panamanian Dry Forests with 15.4% species loss, and finally Talamancan Montane Forests with 15.3% species loss. The variation in estimated species loss between these most affected ecoregions was only 0.2%.

In contrast, the ecoregions in the northwest of the study area and in the Yucatan peninsula had the lowest proportion of species loss, varying from 13.3% in the Yucatan Dry Forests, to 12.7% in the Trans-Mexican Volcanic Belt Pine-Oak Forests, and finally to 12.2% in the Tehuacan Valley Matorral.

These figures indicated that the variation in proportion of estimated species loss between ecoregions was small. This species loss was driven by the change in Effective Habitat Area (or REHA, the proportion of habitat remaining for original species). This means that southern regions lost more of their original species than the northern regions. However, the difference was only 3.4%. Although the net effect of land use change on species loss was greater in the southern regions of Mesoamerica than in the northern regions, all ecoregions seem to have experienced a similar proportion of species loss.

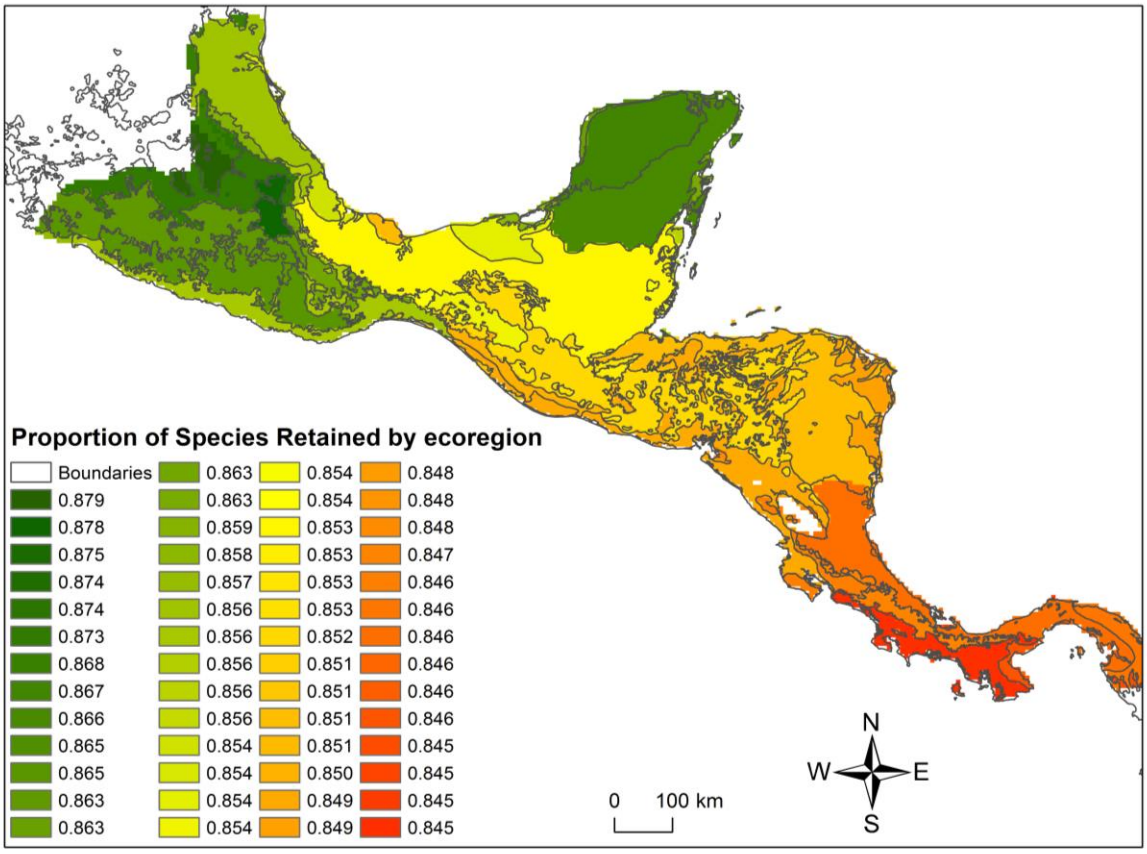


Figure 3.12. Estimated proportion of species retained by ecoregion up to 2000. Estimates are based on the modified SAR method (Allnutt et al. 2008).

3.4 DISCUSSION

3.4.1 Land use change reduces plant habitat in Mesoamerica by 50%

The analysis of the effects of land use change on species diversity indicated that the Mesoamerican hotspot lost more than half of its original habitat prior to 2000. Although the southern ecoregions lost the most habitat, the central and northern ecoregions of the isthmus also experienced high habitat loss. Consequently, predicted species loss prior to the year 2000 across all ecoregions did not show much variation between ecoregions. This is illustrated by the fact that the ecoregion estimated to have had the highest species loss (Isthmian Pacific Moist Forests) lost only 3.3% more than the ecoregion estimated to have had the least loss (Tehuacan Valley Matorral).

Land use change in the form of clear-cutting, selective logging and fire (Ferraz et al. 2003) is a continuous cause of habitat loss and consequent species loss in the region. In this analysis, the concept of Effective Habitat Area (EHA), the proportion of total habitat within a study region that is similar to a focal cell, was

used to estimate the effect of land use change on tree species turnover (beta diversity). In this case, species turnover was modelled using GDM, and the effect of land use change on plant species was directly incorporated through the use of the MSA measure. MSA was used as a measure of habitat condition to differentiate between different types of disturbance. It is a measure of the effect of disturbance on plants, based on a meta-analysis (Alkemade et al. 2009). It is a measure not only of habitat destruction (as in clear cutting) but also of habitat degradation (as in plantation habitats, secondary forests and selective logging). As such, it goes beyond other measures of land use change based only on remaining vegetation cover (Myers et al. 2000), natural intact vegetation (Sloan et al. 2014) or deforestation rates (Brooks et al. 2002). The approach of combining predicted species turnover derived from the GDM model and the MSA measure of habitat condition represents a novel application of the SAR (species area relationship) to estimating diversity loss.

Throughout the region, the EHA values across all grid cells decreased. This was due to the reduction in habitat condition values caused by land use change up to 2000. Referring back to equation (4), the aggregate similarity (sum of similarities) for a focal cell compared to all other cells is the product of the habitat condition (MSA) of the comparison cells and their similarity to the focal cell (s_{ij}). As the similarity or the condition decreases, so does the aggregate similarity and consequently the total effective habitat area (EHA) similar to the focal cell. In this case, the similarity values obtained from the GDM model did not change; only the habitat condition values changed, reflecting the disturbed state as of the year 2000. In **Figure 3.8** and **Figure 3.9**, the overall EHA values decrease as a result of land use change. The proportion of habitat remaining for original species or REHA was calculated as the change in EHA from the assumed undisturbed state to the disturbed state in 2000.

Previous studies have derived varying estimates of remaining habitat for the entire Mesoamerican region, ranging from 14% to almost 54% (**Table 3.3**). This analysis provided estimates for all ecoregions within the Mesoamerican hotspot with an estimated average REHA across all ecoregions of 53.3%. Studies that used data from the 1990s estimated that 20% of primary habitat remained in Mesoamerica at the beginning of the 21st century (**Table 3.3**). Later studies, based on the Global Land Cover 2000 data set (Bartholome et al. 2005), increased that estimate to almost 54%. The average REHA derived in this current analysis fell within this range at 53.3%. It is clear that studies that derived estimates using the Global Land Cover 2000 data set all show estimates similar to each other. Unlike the present study, previous investigations focused only on vegetation cover and did not incorporate species turnover or habitat condition in their estimates. However, this analysis showed similar results despite the incorporation of these factors. The effect of incorporating both species turnover and MSA in this study seems to be the same as only taking into account vegetation cover.

The latest investigation estimated remaining natural intact vegetation (NIV) at 14.1% (Sloan et al. 2014). These authors did not take into account species turnover, however. In contrast, the current study is based on

EHA, an areal estimate that included similarity between sites or cells (species turnover). Furthermore, the estimates presented in this analysis also incorporated the MSA measure of habitat disturbance, which is a proportional measure of the effect of land use change on plants. Other factors that may have contributed to the low estimates produced by Sloan et al. (*ibid*) are: (1) their approach focused only on assumed natural vegetation, therefore it excluded cover from timber and crop plantations; (2) they used more recent global GlobCover 2009 satellite-image classification data; and (3) their method removed other "disturbed" areas such as small forest fragments, burned sites, settlements, and areas adjacent to roads. In the case of timber plantation areas, for example, the current analysis assigned an MSA value of 0.2 rather than exclude such areas altogether. In addition, unlike Sloan et al. (*ibid*), the current analysis also did not take into account the effect of infrastructure, such as roads. Such methodological differences are likely to drive the large disparity in estimates of remaining habitat between these two studies.

The fact that the change in EHA (or REHA) estimates presented here were so similar to previous estimates that only took into account vegetation cover could be partially related to the model predictions. The effect of varying similarity (s_{ij}) values in calculating change in EHA was to alter the area for which habitat condition (c_j) matters to the change in EHA from the pristine condition to the condition after land use change (see equation 3). In this analysis, the similarity values varied between 0 and 0.2 because the GDM dissimilarity values varied between 0.8 and 1 (**see Chapter 2**). Low similarity values reduced the impact of the habitat condition of other grid cells on the overall change in EHA of a focal cell. If a focal cell was totally different from all other grid cells (S_{ij}=0), then the change in EHA would be determined solely by the change in habitat condition of that focal cell because no other cells have similar habitat. Conversely, if a focal cell was totally identical to all other grid cells, then its change in EHA would be dependent on the change in habitat condition of those cells. Therefore, in this context, the effect of habitat loss or degradation (as measured by MSA) varied widely depending on the similarity of each particular grid cell to all other cells.

It is important to note that the method implemented in this analysis may not be directly comparable to the methods used in the other studies. EHA is a measure of similar habitat and REHA shows the change in area of similar habitat. The studies cited here do not take into account spatial species turnover; therefore, similar values for remaining habitat may be purely coincidental. Since the method implemented here incorporates spatial species turnover (beta diversity), and consequently the identity of species, it can be argued that this method provides a more refined way of representing diversity. However, the purpose here is not to promote one method over the other (Brooks et al. 2002), but to demonstrate how an understanding of species turnover patterns can provide further insight about centres of diversity and complement analyses of diversity loss.

Table 3.3 Estimates of natural vegetation area as percentage of original vegetation area, from different studies.

Estimates from literature	Percentage (%)	Baseline data dates[a]
Myers et al. (2000): *Primary vegetation*	20	1992-1997
Mittermeier et al. (2004): *Primary habitat*	20	1990-1998
Hoekstra et al. (2005): *Undisturbed habitat*	53.8	2000
Schmitt et al. (2009): *Relatively natural forest cover*	52.4	2000
Sloan et al. (2014): *Natural intact vegetation*	14.1	2009-2012
Present Study: *Effective habitat area*	53.3 (average across ecoregions)	2000

[a] Dates obtained from Myers et al. (2000) and Sloan et al. (2014).

3.4.2 Land use change drives high species loss across all ecoregions in Mesoamerica

In this analysis, estimated habitat loss varied from 40.6% to 49.1% with a mean loss of 46.7% across all ecoregions. Using the SAR, a total of 14.3% tree species were estimated to be extinct from the region by the year 2000 due to the loss in effective habitat area. Despite having lost almost half of the EHA in the region, less than 15% of originally occurring species were predicted to have been eliminated. Previous estimates that combined the rate of habitat loss, the SAR, and survivorship curves, estimated that the clearing of half the humid forests eliminated 15% of the species they contained (Pimm and Raven 2000). However, these figures could be an artefact of the non-linearity of the SAR curve (Pimm and Raven 2000). A study that applied deforestation rates in using the SAR estimated that 33.1% of species should have been threatened or extinct in Mesoamerica by 2000 (Brooks et al. 2002). The same authors noted, however, that Red List data from 1998 indicated that only 19% of species in the region were threatened or extinct (Brooks et al. 2002). Within this context, the value obtained from utilizing the EHA method was at the lower end of estimates since it was based on the loss of similar habitat. By including beta diversity, this analysis accounted for the fact that habitat destruction could "stamp out" species in one area (Pimm and Raven 2000) but these species could have remained in other areas.

Several issues with the use of the SAR have been highlighted in the literature. The wide incongruity between SAR-predicted extinctions and actual recorded extinctions has been attributed to "extinction debt", that is, there may be a time lag between habitat loss and the extinction of a species' last viable population (He and Hubbell 2011). Recent analyses have shown that the SAR always overestimates species loss from habitat loss because it assumes the same underlying sampling method for both the forward species accumulation curve and the backward SAR (used to estimate species extinction). In reality, the assumption is unjustified because the area that must be added to find the first individual of a species (species accumulation curve) is much smaller than the area that must be removed to eliminate the last individual of a species (He and Hubbell

2011; Rahbek and Colwell 2011). This means that the backward SAR predicts a higher rate of species extinctions from habitat loss than is characteristic of real communities, and hence overestimates species loss.

The EAR (endemic area relationship) has been proposed as the curve that correctly describes species extinction from habitat loss and hence could provide better estimates (Rahbek and Colwell 2011). It is possible that the method presented here also offered a better way of approximating species loss due to habitat loss because it referred to loss of similar habitat based on similarity between areas (beta diversity), which inherently accounted for species identity. Beta diversity is driven to a large extent by endemism; therefore, in this method proportion of species loss depended on whether the entire habitat for any given species had been lost. However, this study did not implement the EAR method, so comparisons are not possible. Further research could investigate whether this modified-SAR method and the EAR provide similar results, or whether these relationships based on the current plant data converge into a universal relationship, as has been shown for vertebrates taxa (Storch et al. 2012).

A related issue is whether the 'z' exponent should have been applied at all in the calculations implemented in the modified-SAR method presented here. The slope variable 'z' measures the rate of species increase for the SAR power law function, and has been suggested to be directly proportional to beta diversity (Mashayekhi et al. 2014). If 'z' represents beta diversity and they both essentially measure the same parameter, then this analysis already accounted for this parameter via similarity among cells. If no exponent had been applied then the proportion of species retained would have been equivalent only to the change in EHA. However, this would have led to an increase in estimated species loss.

Finally, this analysis did not apply different 'z' values to forested and non-forested areas. Studies have shown varying values between forested and non-forested areas (Drakare et al. 2006), and for comparisons based on nested situations versus island situations (Dirzo and Raven 2003). Here, the REHA was based on remaining forested and non-forested habitat, so the same exponent was applied universally. Other studies identify higher 'z' values for forested habitat ($z=0.48$) than for non-forested habitat ($z=0.23$) (Drakare et al. 2006). It may have been beneficial to make such a distinction and applied a higher exponent to pixels that remained forested after land use change and a lower exponent to pixels that became non-forested. However, in the context of the present analysis, this distinction could only have been made by referring back to the habitat condition value (MSA), but this seems somewhat tautological.

The discrepancies among methods for estimating species loss from habitat loss and resulting variation in species extinction estimates does not negate the fact that habitat loss is a driver of species extinctions, and continues to threaten the long term persistence of biodiversity (Koh and Ghazoul 2010). On the contrary, these studies highlight the gaps in our knowledge and that it is prudent to apply the precautionary principle. This means putting in place measures to protect and sustainably use any remaining habitat.

3.4.3 Conservation implications

This chapter provides an overview of the intra-regional variation in habitat loss and species loss within the Mesoamerican region. Most of the region lost half of its originally occurring habitat. This is not surprising as the Mesoamerican region is a rapidly developing region with competing development priorities. This analysis provides a snapshot of the situation at the year 2000. Fifteen years have passed since, suggesting that the situation may have deteriorated if land use change continued in a similar pattern. Using newer data on forest cover change, such as that developed by Hansen et al. (2013), would help update the estimates presented here. Unfortunately, this updated data was not available for use at the time of this analysis.

Nevertheless, the current results emphasized that the southern ecoregions in Costa Rica and Panama lost much of their original habitat and species. Importantly, this variation cannot be attributed solely to the narrower area of these countries (compared to the rest of the region) since the current method accounted for the effect of area (see **Methods**). It is crucial for the region's Pine-Oak Forests, Montane Forests and Dry Forests, most of which show low REHA, to be conserved. Only 1.7% of the region's Dry Forests remain and are restricted to the Pacific coasts of Mexico, Nicaragua, Costa Rica and Panama (Griscom and Ashton 2011). The Dry Forests in Mexico provide a good opportunity for conservation efforts since they were the exception and showed high REHA but low protection status (see **Chapter 4**).

Importantly, the results presented here have occurred in other studies. This analysis and the most recent studies (**Table 3.4**) refer to estimated habitat loss and diversity loss in ecoregions up to 2000 (Newbold et al. 2015), and up to 2012 in the case of Sloan et al. (2014). The earlier studies predict changes in ecoregion species richness based on climate change scenarios (Golicher et al. 2012), or identify priority areas based on species richness and climate change (Anderson et al. 2008). Despite the difference in methods used here and those of other investigations, such disparate methods revealed similar patterns of conservation priority. For example, my results showed the same ecoregions had the lowest REHA (first quartile) as those that had low Natural Intact Vegetation or NIV (Sloan et al. 2014) (see **Table 3.4**). Some notable differences between this and their study include Balsas Dry Forests, Talamancan Montane Forests and Choco-Darien Moist Forests. In the case of Balsas Dry Forests, this study found that ecoregion to have high REHA in 2000, but they found it to have the lowest NIV in 2012. One possible reason for this discrepancy is that they used more updated land use data (**Table 3.3**). This just highlights the fact that habitat loss continued to affect this ecoregions after 2000. In the case of the other two ecoregions, the current analysis found below average REHA but they found high NIV. Again, their updated land use data could account for this discrepancy, and could be due to increased conservation efforts in these ecoregions after the year 2000. Nevertheless, I propose that the ecoregions in Table 3.4 provide a good starting point for focusing conservation efforts within the Mesoamerican hotspot.

Table 3.4. Priority ecoregions within the Mesoamerican hotspot, as defined by four recent studies and this one. Ecoregions are ranked in descending order by the number of studies that identified them as a priority.

Ecoregions	Anderson et al. (2008)	Golicher et al. (2011)	Sloan et al. (2014)	Newbold et al. (2015)	Present study
Isthmian-Pacific Moist Forests		x	x	x	x
Panamanian Dry Forest			x	x	x
Isthmian-Atlantic Moist Forests	x	x			x
Talamancan Montane Forests	x			x	x
Costa Rican Seasonal Moist Forest			x		x
Central American Atlantic Moist Forests	x	x			
Chiapas Depression Dry Forest			x	x	
Miskito Pine Forests	x				x
Central American Dry Forest			x		
Montagua Valley Thornscrub			x		
Pantanos de Centla			x		
Balsas Dry Forests			x		
Southern Pacific Dry Forests				x	
Veracruz Moist Forests				x	
Peten-Veracruz Moist Forests				x	
Central American Pine-Oak Forests				x	
Central American Montane Forests				x	
Eastern Panamanian Montane Forests					x
Choco-Darien Moist Forests					x
Southern Mesoamerican Pacific Mangroves					x

3.5 CONCLUSION

This study represents the only instance within the region where beta diversity and the SAR were combined to estimate loss of originally occurring habitat and species. The results of this analysis showed that the entire Mesoamerican region had lost almost half of its original habitat until 2000, condemning close to 15% of originally occurring species to extinction. Habitat loss and species extinction were estimated highest in the southern ecoregions, encompassed in Costa Rica and Panama. Many studies investigate the effect of land use change and habitat loss on species richness. This analysis highlights the additional insight that an understanding of species spatial turnover (beta diversity) patterns can provide in identifying centres of diversity and in prioritizing areas for conservation. This study focused on tree data because trees represent the primary productive sector in ecosystems. It is important to remember, however, that all taxa play a role in ecosystem resilience, and it is equally crucial to quantify the effect of land use change on their assemblages. Furthermore, although this study focused on tree assemblages at the ecoregion level, it must be noted that species-level conservation is also an important part of ecosystem conservation. The loss of native

species in exchange for non-native timber plantations, for example, may require species-focused conservation efforts. The results presented here serve as a cautionary tale about the effect of extensive habitat loss that has occurred in the Mesoamerican hotspot. As developing nations, Mesoamerican countries must balance poverty alleviation and development needs with the conservation efforts necessary to maintain the ecosystems that underpin human well-being. Initiatives such as the Central American Commission on Environment and Development have been developed to address these challenges with a more concerted effort, and provide hope that the region can take substantial steps towards a more sustainable future.

3.6 REFERENCES

Alkemade, R., Oorschot, M., Miles, L., Nellemann, C., Bakkenes, M. and ten Brink, B., 2009. GLOBIO3: A Framework to Investigate Options for Reducing Global Terrestrial Biodiversity Loss. *Ecosystems*, *12* (3), 374–390.

Allnutt, T. F., Ferrier, S., Manion, G., Powell, G. V. N., Ricketts, T. H., Fisher, B. L., Harper, G. J., Irwin, M. E., Kremen, C., Labat, J.-N., Lees, D. C., Pearce, T. a. and Rakotondrainibe, F., 2008. A method for quantifying biodiversity loss and its application to a 50-year record of deforestation across Madagascar. *Conservation Letters*, *1* (4), 173–181.

Anderson, E. R., Cherrington, E. A., Flores, A. I., Perez, J. B., Carrillo, R. and Sempris, E., 2008. *Potential Impacts of Climate Change on Biodiversity in Central America, Mexico, and the Dominican Republic*. *Ecosystems*. Panama City, Panama

Anderson, E. R., Cherrington, E. A., Tremblay-Boyer, L., Flores, A. I. and Sempris, E., 2008. Identifying Critical Areas for Conservation : Biodiversity and Climate Change in Central America , Mexico , and the Dominican Republic. *Biodiversity*, *9* (3 & 4), 89–99.

Anderson, M. J., Crist, T. O., Chase, J. M., Vellend, M., Inouye, B. D., Freestone, A. L., Sanders, N. J., Cornell, H. V., Comita, L. S., Davies, K. F., Harrison, S. P., Kraft, N. J. B., Stegen, J. C. and Swenson, N. G., 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology Letters*, *14* , 19–28.

Arrhenius, O., 1921. Species and Area. *The Journal of Ecology*, *9* (1), 95–99.

Bartholome, E., Belward, A., Beuchle, R., Eva, H., Fritz, S., Hartley, A., Mayaux, P. and Stibig, H. J., 2005. Global land cover for the year 2000. *International Journal of Remote Sensing*, *26* (9), 1959–1977.

Benito, B. M., Cayuela, L. and Albuquerque, F. S., 2013. The impact of modelling choices in the predictive performance of richness maps derived from species-distribution models: guidelines to build better diversity models. *Methods in Ecology and Evolution*, *4* (4), 327–335.

Broennimann, O., Thuiller, W., Hughes, G., Midgley, G. F., Alkemade, J. M. R. and Guisan, A., 2006. Do geographic distribution, niche property and life form explain plants' vulnerability to global change? *Global Change Biology*, *12* (6), 1079–1093.

Brooks, T. M., Mittermeier, R. A., da Fonseca, G. A. B., Gerlach, J., Hoffmann, M., Lamoreux, J. F., Mittermeier, C. G., Pilgrim, J. D. and Rodrigues, A. S. L., 2006. Global Biodiversity Conservation Priorities. *Science*, *313* (July), 58–61.

Brooks, T. M., Mittermeier, R. a., Mittermeier, C. G., da Fonseca, G. a. B., Rylands, A. B., Konstant, W. R., Flick, P., Pilgrim, J., Oldfield, S., Magin, G. and Hilton-Taylor, C., 2002. Habitat Loss and Extinction in the Hotspots of Biodiversity. *Conservation Biology*, *16* (4), 909–923.

Cayuela, L., Granzow-de la Cerda, Í., Albuquerque, F. S. and Golicher, D. J., 2012. Taxonstand: An r package for species names standardisation in vegetation databases. *Methods in Ecology and Evolution*, *3* (6), 1078–1083.

Conservation International, 2004. *CEPF Ecosystem Profile: Northern Region Of The Mesomaerica Biodiversity Hotspot*. *Analysis*. Washington, D.C.

Dirzo, R. and Raven, P. H., 2003. Global State of Biodiversity and Loss. *Annual Review of Environment and Resources*, *28* (1), 137–167.

Drakare, S., Lennon, J. J. and Hillebrand, H., 2006. The imprint of the geographical, evolutionary and ecological context on species-area relationships. *Ecology Letters*, *9* (2), 215–27.

Fairbanks, D., Reyers, B. and van Jaarsveld, A. S., 2001. Species and environment representation: selecting reserves for the retention of avian diversity in KwaZulu-Natal, South Africa. *Biological Conservation*, *98* (3), 365–379.

Feeley, K. J. and Silman, M. R., 2011. The data void in modeling current and future distributions of tropical species. *Global Change Biology*, *17* (1), 626–630.

Feeley, K. and Silman, M., 2010. Land‐use and climate change effects on population size and extinction risk of Andean plants. *Global Change Biology*, *16* , 3215–3222.

Ferraz, G., Russell, G. J., Stouffer, P. C., Bierregaard, R. O., Pimm, S. L. and Lovejoy, T. E., 2003. Rates of species loss from Amazonian forest fragments. *Proceedings of the National Academy of Sciences of the United States of America*, *100* (24), 14069–73.

Ferrier, S., Drielsma, M., Manion, G. and Watson, G., 2002. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales . II . Community-level modelling. *Biodiversity and Conservation*, *11* , 2309–2338.

Ferrier, S. and Guisan, A., 2006. Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, *43* (3), 393–404.

Ferrier, S., Powell, G. V. N., Richardson, K. S., Manion, G., M, J., Allnutt, T. F., Cameron, S. E., Mantle, K., Burgess, N. D., Faith, D. P., Lamoreux, J. F., Kier, G., Hijmans, R. J., Funk, V. A., Cassis, G. A., Fisher, L., Flemons, P., Lees, D., Lovett, J. C., Rompaey, R. S. A. R. Van, Overton, J. M., Allnutt, F., Fisher, B. L., Lovett, J. O. N. C. and Rompaey, R. S. A. R. V. A. N., 2004. Mapping More of Terrestrial Global Biodiversity for Conservation Assessment. *BioScience*, *54* (12), 1101–1109.

Ferrier, S., 2002. Mapping Spatial Pattern in Biodiversity for Regional Conservation Planning : Where to from Here ? *Systematic Biology*, *51* (2), 331–363.

Funk, S. M. and Fa, J. E., 2010.b Ecoregion prioritization suggests an armoury not a silver bullet for conservation planning. *PloS One*, *5* (1), 1–7.

Golicher, D. J., Cayuela, L. and Newton, A. C., 2012. Effects of Climate Change on the Potential Species Richness of Mesoamerican Forests. *Biotropica*, *44* (3), 284–293.

Griscom, H. P. and Ashton, M. S., 2011. Restoration of dry tropical forests in Central America: A review of pattern and process. *Forest Ecology and Management*, *261* (10), 1564–1579.

Hansen, M. C., Potapov, P. V, Moore, R., Hancher, M., Turubanova, S. a, Tyukavina, A., Thau, D., Stehman, S. V, Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O. and Townshend, J. R. G., 2013. High-resolution global maps of 21st-century forest cover change. *Science*, *342* (6160), 850–853.

He, F. and Hubbell, S. P., 2011. Species-area relationships always overestimate extinction rates from habitat loss. *Nature*, *473* (7347), 368–71.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. and Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25* (15), 1965–1978.

Jennings, M. D., Hoekstra, J., Higgins, J. and Boucher, T., 2008. A comparative measure of biodiversity based on species composition. *Biodiversity and Conservation*, *17* (4), 833–840.

Jetz, W., Wilcove, D. S. and Dobson, A. P., 2007. Projected impacts of climate and land-use change on the global diversity of birds. *PLoS Biology*, *5* (6), e157.

Koh, L. P. and Ghazoul, J., 2010. A matrix-calibrated species-area model for predicting biodiversity losses due to land-use change. *Conservation Biology*, *24* (4), 994–1001.

Koleff, P. and Gaston, K. J., 2002. The relationships between local and regional species richness and spatial turnover. *Global Ecology and Biogeography*, *11* (5), 363–375.

Leathwick, J. R., Moilanen, a., Ferrier, S. and Julian, K., 2010. Complementarity-based conservation prioritization using a community classification, and its application to riverine ecosystems. *Biological Conservation*, *143* (4), 984–991.

Londoño-Murcia, M. C., Tellez-Valdés, O. and Sánchez-Cordero, V., 2010. Environmental heterogeneity of World Wildlife Fund for Nature ecoregions and implications for conservation in Neotropical biodiversity hotspots. *Environmental Conservation*, *37* (02), 116–127.

Margules, C. R. and Pressey, R. L., 2000. Systematic conservation planning. *Nature*, *405* (6783), 243–53.

Mashayekhi, M., MacPherson, B. and Gras, R., 2014. Species–area relationship and a tentative interpretation of the function coefficients in an ecosystem simulation. *Ecological Complexity*, *19* , 84–95.

Miller, K., Chang, E. and Johnson, N., 2001. *Defining common ground for the Mesoamerican biological corridor. Management*. Washington, D.C.: World Resources Institute

Morawetz, W. and Raedig, C., 2007. Angiosperm Biodiversity , Endemism and Conservation in the Neotropics. *Taxon*, *56* (4), 1245–1254.

Mu, Q., Zhao, M. and Running, S. W., 2011. Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sensing of Environment*, *115* (8), 1781–1800.

Muller, R., Nowicki, C., Barthlott, W. and Ibisch, P. L., 2003. Biodiversity and endemism mapping as a tool for regional conservation planning – case study of the Pleurothallidinae (Orchidaceae) of the Andean rain forests in Bolivia. *Biodiversity and Conservation*, *12* , 2005–2024.

Myers, N., Mittermeier, R. a, Mittermeier, C. G., da Fonseca, G. a and Kent, J., 2000. Biodiversity hotspots for conservation priorities. *Nature*, *403* (6772), 853–858.

Newbold, T., Hudson, L. N., Hill, S. L. L., Contu, S., Lysenko, I., Senior, R. a., Börger, L., Bennett, D. J., Choimes, A., Collen, B., Day, J., De Palma, A., Díaz, S., Echeverria-Londoño, S., Edgar, M. J.,

Feldman, A., Garon, M., Harrison, M. L. K., Alhusseini, T., Ingram, D. J., Itescu, Y., Kattge, J., Kemp, V., Kirkpatrick, L., Kleyer, M., Correia, D. L. P., Martin, C. D., Meiri, S., Novosolov, M., Pan, Y., Phillips, H. R. P., Purves, D. W., Robinson, A., Simpson, J., Tuck, S. L., Weiher, E., White, H. J., Ewers, R. M., Mace, G. M., Scharlemann, J. P. W. and Purvis, A., 2015. Global effects of land use on local terrestrial biodiversity. *Nature*, *520* (7545), 45–50.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'amico, J. a., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P. and Kassem, K. R., 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience*, *51* (11), 933-938.

Pimm, S. L. and Raven, P., 2000. Extinction by numbers. *Nature*, *403* , 843–845.

Pressey, R. L., Cabeza, M., Watts, M. E., Cowling, R. M. and Wilson, K. A., 2007. Conservation planning in a changing world. *Trends in Ecology & Evolution*, *22* (11), 583–592.

Rahbek, C. and Colwell, R. K., 2011. Species loss revisited. *Nature*, *473* , 299–289.

Sloan, S., Jenkins, C. N., Joppa, L. N., Gaveau, D. L. A. and Laurance, W. F., 2014. Remaining natural vegetation in the global biodiversity hotspots. *Biological Conservation*, *177* , 12–24.

Storch, D., Keil, P. and Jetz, W., 2012. Universal species-area and endemics-area relationships at continental scales. *Nature*, *488* (7409), 78–81.

Tuomisto, H. and Ruokolainen, K., 2006. Analyzing or explaining beta diversity? Understanding the targets of different methods of analysis. *Ecology*, *87* (11), 2697–2708.

Tuomisto, H., 2010.a A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography*, *33* (1), 2–22.

Tuomisto, H., 2010.b A diversity of beta diversities: straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. *Ecography*, *33* (1), 23–45.

Whittaker, R. H., 1972. Evolution and Measurement of Species Diversity. *Taxon*, *21* (2/3), 213–251.

Wu, F., Yang, X. J. and Yang, J. X., 2010. Additive diversity partitioning as a guide to regional montane reserve design in Asia: an example from Yunnan Province, China. *Diversity and Distributions*, *16* (6), 1022–1033.

Xiao-jun, L., Boucher, T. and Sayre, R., 1999. A conservation assessment of Central American vegetation and ecoregions. *Journal of Natural Resources*, *4* (4), 300–306.

4 CHAPTER 4

Assessing Representativeness of the Protected Area System

4.1 INTRODUCTION

Habitat and biodiversity loss are phenomena of an ever-expanding human civilization. Having realized the inter-connectedness of human society with its natural environment, much effort has been placed in recent decades on slowing habitat and diversity loss (Gaston et al. 2008). One of the major steps towards this goal has been the establishment of protected natural areas (Gaston et al. 2008). Such areas are established for a diversity of goals including protecting habitat, keystone species or suites of species (Margules and Pressey 2000), but they all require significant financial and management investments. As such, prioritizing the location and effectiveness of such areas is a crucial factor in their success.

Conservation planning is the process of locating, configuring and maintaining areas to promote the persistence of biodiversity (Pressey et al. 2007). Research in this area has led to the development of systematic processes or methodologies aimed at achieving conservation objectives. Since it was first conceived, systematic conservation planning has informed policy, legislation and on-the-ground conservation actions (Pressey et al. 2007). Major conservation organization such as The Nature Conservancy have utilized systematic conservation planning processes to focus funding (Pressey et al. 2007). Such processes include, among other stages, setting quantitative targets for representation of species, vegetation types, or other features of interest (Margules and Pressey 2000). For example, one process developed for marine protected areas in Victoria, Australia calls for protected areas to be comprehensive, adequate, and representative (CAR) (Wescott 2006). One commonality among different methodologies is that they call for protected areas to hold representative samples of the habitats or species they aim to protect. Measures employed to assess representativeness include species richness, coverage of threatened or endangered species, or percentage coverage of habitats at different scales e.g. ecoregions (Jenkins and Joppa 2009; A. Rodrigues et al. 2004; Soutullo et al. 2007; Wilson et al. 2007).

A recent addition to the criteria of representativeness has been that protected areas be representative not only of present species or habitats but also of the ecological processes that maintain biodiversity. Such processes can be planned for by employing approaches such as variable representation targets, moveable conservation areas, specific design criteria, and spatial catalysts (Pressey et al. 2007). Spatial catalysts refer to features that can be used as surrogates for ecological processes. Spatial catalysts can include edaphic factors such as soil interfaces, which mark the boundaries between soils that drive large differences in species composition between areas (Pressey et al. 2007). It seems logical, then, that protecting areas that encompass changes in species composition will also encompass important spatial catalysts and thus processes that drive species diversification.

The concept of beta diversity can be used in conservation planning as an indicator of the location of factors associated with long term processes and species persistence (Wiersma and Urban 2005). Beta diversity provides a means of assessing which areas encompass changes in species composition. In this investigation, beta diversity is defined as a measure of the amount of variation in species composition across sites within a specific region (Tuomisto and Ruokolainen 2006). Beta diversity can be expressed using various dissimilarity indices that quantify species spatial turnover or the difference in species composition between any pair of sites (Koleff et al. 2003). This study applies this concept of beta diversity to assess the representativeness of protected areas in Mesoamerica in representing tree species diversity. Evaluating the representativeness of protected areas can serve as a basis for determining gaps in the extent of biodiversity protection and inform decision-makers about conservation priorities (Chape et al. 2005).

The objective of this chapter is to assess the representativeness of the regional protected areas system in relation to species composition turnover patterns. This will be achieved by addressing the following questions:

1. How representative of tree diversity is the PA system as a whole?
2. What ecoregions are best or least represented in the PA system?
3. What environmental factors are associated with level of protection within the protected area system?

4.2 METHODS

In order to assess the representativeness of the regional protected areas network, the predictors of compositional dissimilarity derived from a GDM model are combined with information on protected area coverage across the region to estimate the proportion of tree diversity captured by the reserve network. The process is similar to that described in assessing the proportion of species retained after land use change (see Figure 3.1, Chapter 3). In this case, instead of using MSA as a measure of habitat condition, the protected status of each 0.1 degree grid cell is used to estimate the proportion of similar habitat represented within the reserve system (**Figure 4.1**). The analysis was implemented using the software package Muru: Pathways to Resilience developed by Tom Harwood at CSIRO. The steps in the process are described in more detail below.

(i) Model of compositional dissimilarity:

The predicted dissimilarity values for all 0.1 degree grid cells in the study area obtained from the eleven predictor GDM model described in **Chapter 2** are used to conduct this assessment. For the purposed of this chapter, I calculate what I term 'dissimilarity values' for each grid cell. These values are in fact the reciprocal of EHA (1/EHA), and are then multiplied by 100 to allow ease of comparison with the proportion of habitat represented. For a visual representation of the spatial pattern of compositional similarity as represented by EHA, see **Figure 2.2**.

(ii) Measure of protected status:

A raster grid was derived to represent the protected status of each cell in the study area. This was derived by converting a polygon layer of protected areas to a raster grid using the 'Polygon to Raster' tool in ArcGIS 10. The resolution of the resulting raster was $1km^2$. This $1km^2$ raster was then reclassified to 0 (not protected) and 1 (protected), then re-sampled to $10km^2$ (the 0.1 degree resolution used for this study) by averaging cell values. This means that the resulting $10km^2$ resolution raster was a continuous raster where each $10km^2$ grid cell had a value ranging from 0 to 1, representing the proportion of the cell that was under protected status. This raster was then used for this analysis in a similar way that the continuous MSA grid was used in **Chapter 3**. The base data used to conduct this classification was from the WPDA 2010 (World Database on Protected Areas, IUCN and UNEP 2010, www.protectedplanet.net). It is important to note that the term "protected" refers to all IUCN categories of protected status including those which may allow extraction activities, such as forest reserves. That is, this analysis assumes that all protected areas have an equal level of protection, and does not distinguish between levels of effectiveness in protecting biodiversity.

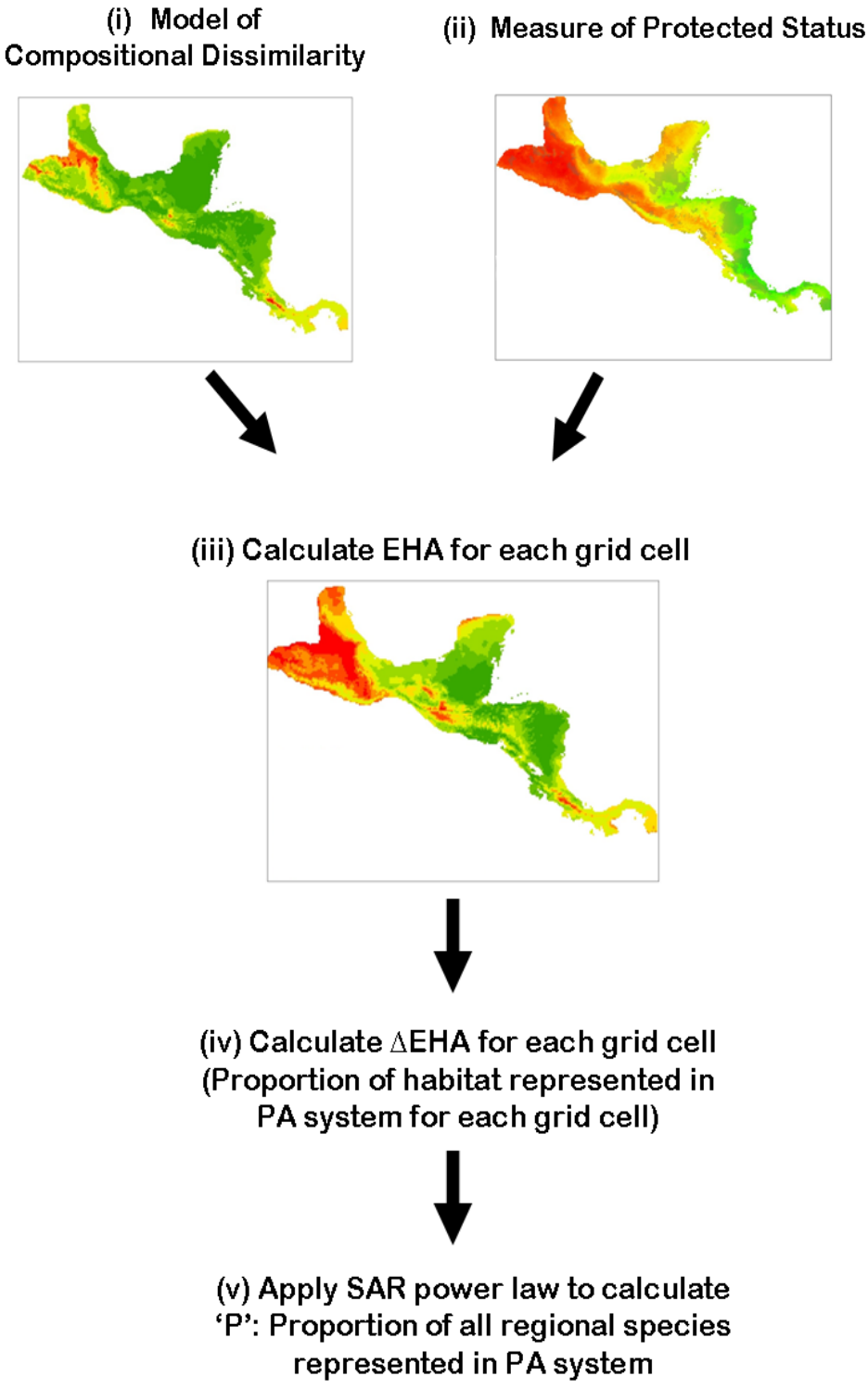


Figure 4.1 Framework for assessing representativeness of Protected Area system.

(iii) Calculate EHA for each grid cell:

This analysis follows an approach used in previous studies (Allnutt et al. 2008; Ferrier et al. 2004, 2012; Prober et al. 2012; Williams et al. 2010). Using the information on pairwise dissimilarity and protected status of each 0.1 degree grid cell, the "effective habitat area" (EHA) was calculated for each cell. EHA is the total area of habitat predicted to be similar in biological composition to the cell of interest (Williams et al. 2010). For any given cell, its total EHA was calculated as:

$$EHA_i = \sum_{j=1}^{n} s_{ij} h_j \qquad \text{(equation 5)}$$

where 'i' is the focal cell, s_{ij} is the similarity (1-dissimilarity) between cells 'i' and 'j' and h_j is the protected status of cell 'j' (Williams et al. 2010). If the similarity between these cells is 1 and the protected status of cell 'j' is also 1, this means that one entire cell out of all 'n' cells has habitat that is totally compositionally similar to cell 'i'. If the similarity or protected status of cell 'j' is decreased, so will its contribution to the effective area of similar habitat to cell 'i' (Williams et al. 2010). In this way, the total EHA of cell 'i' was calculated as the sum of the contributions in similarity and protected status of all other 'n' cells in the study region.

(iv) Calculate ΔEHA for each grid cell:

Once the EHA for a cell was calculated, it was expressed as a proportion of the EHA obtained if all cells in the study region were assumed to be protected, i.e, have an 'h' value of 1 (Williams et al. 2010). In this way, ΔEHA represents the change in EHA when going from a state where all cells in the region are protected to a state where only a proportion of cells are protected. In the former case, since all cells are assumed to be protected, the EHA is simply the sum of similarities ($\sum s_{ij}$), and ΔEHA_i can be calculated as:

$$\Delta EHA_i = \frac{\sum_{j=1}^{n} s_{ij} h_j}{\sum_{j=1}^{n} s_{ij}} \qquad \text{(equation 6)}$$

Implementing this approach, produced a continuous grid in which the value of each cell (ΔEHA_i) is the proportion of the total area of cells with a similar environment, and therefore similar species composition, included in the reserve system (Williams et al. 2010). This is referred to in the results section as the proportion of habitat represented in the PA system for each grid cell. Once ΔEHA was calculated for each cell, the results were aggregated to the ecoregion scale to get a mean proportion of habitat represented in the PA system for each ecoregion. The mean proportion of habitat protected was also calculated for each ecoregion, after accounting for habitat loss up to the year 2009. Calculations were done in R Version 2.13.2.

(v) Apply SAR power law to calculate 'P', proportion of all regional species represented in PA system

The ΔEHA was calculated across all cells in the study region. Then, the SAR power law was applied to estimate the proportion of actual species that are represented in the PA system across the entire region. This was done following the method outlined in Allnutt et al. 2008, and described in detail in **Chapter 3**.

4.3 RESULTS

4.3.1 Protected areas protect up to 32% of tree habitat, representing 67.7% of original tree species

Results for both the four-predictor model and the eleven-predictor model quantified the amount of tree habitat represented within protected areas at below 50%. The eleven-predictor model predicted a mean proportion of tree habitat within protected areas of 22.4%, a maximum of 32.4% and a minimum of 14.1% across all grid cells (**Figure 4.2**). The four-predictor model predicted a mean proportion of habitat represented of 22.1%, a maximum of 30.0% and a minimum of 14.7% across all grid cells. However, the spatial pattern of habitat representation across the region was similar for both models. Since both models yielded similar outputs, hereafter, results are discussed in relation to the eleven predictor model. The areas with the lowest proportion of habitat representation occurred in the west of the study region (**Figure 4.2, A**), in the Yucatan peninsula (**Figure 4.2**, **B**) and along the Pacific coast of the isthmus (**Figure 4.2, C**). The areas with the highest proportion of habitat represented were along the Atlantic coast of the study region extending along the coasts of Belize (**Figure 4.2, D**), Honduras and Nicaragua (**Figure 4.2, E**), and including most of Costa Rica and Panama (**Figure 4.2, F**).

In the west region of the study area, the low proportion of habitat representation in reserves was similar across all grid cells (**Figure 4.2, A**). This could indicate that species turnover in these areas was high, but occurred at a similar rate across these areas. As a result, reserves across any of these areas would have a similar localized effect. That is, any areas outside of reserves in this region were likely to be poorly protected because their flora was likely to be dissimilar to that in a reserve, even though they were geographically close to that reserve. Hence, these areas showed a similar low proportion of habitat representation in reserves.

The other areas of low habitat representation in the Yucatan peninsula and along the Pacific coast showed a less consistent spatial pattern (**Figure 4.2**, **B & C**). This could indicate that there was high variation in rate of species turnover in these areas. Species turnover could have occurred gradually across large distances due to gradual changes in climatic conditions or occurred rapidly across short distances due to steeper climatic gradients. The effect of reserves in these areas would thus be highly variable depending on the rate of species turnover at the site of the reserve. This would lead to high variation in the proportion of habitat

represented. This is evidenced in these specific areas that show grid cells having low habitat representation interspersed with grid cells of intermediate representation

The areas with the highest proportion of reserve representation (**Figure 4.2**, **D**-**F**) showed similar values across all grid cells in these areas. This could indicate that, in these areas, species turnover occurred gradually over large distances due to gradual changes in climatic gradients. Reserves in these areas would have a broad scale effect, where habitats outside of reserves were more likely to be represented within reserves as their flora would be more likely to be similar to that in reserves.

These results for the proportion of tree habitat represented were extended (by applying the SAR power law) to estimate the proportion of assumed original species represented in the PA system across the entire region. The eleven-predictor model estimated that 67.7% of original tree species were represented in the current reserve network.

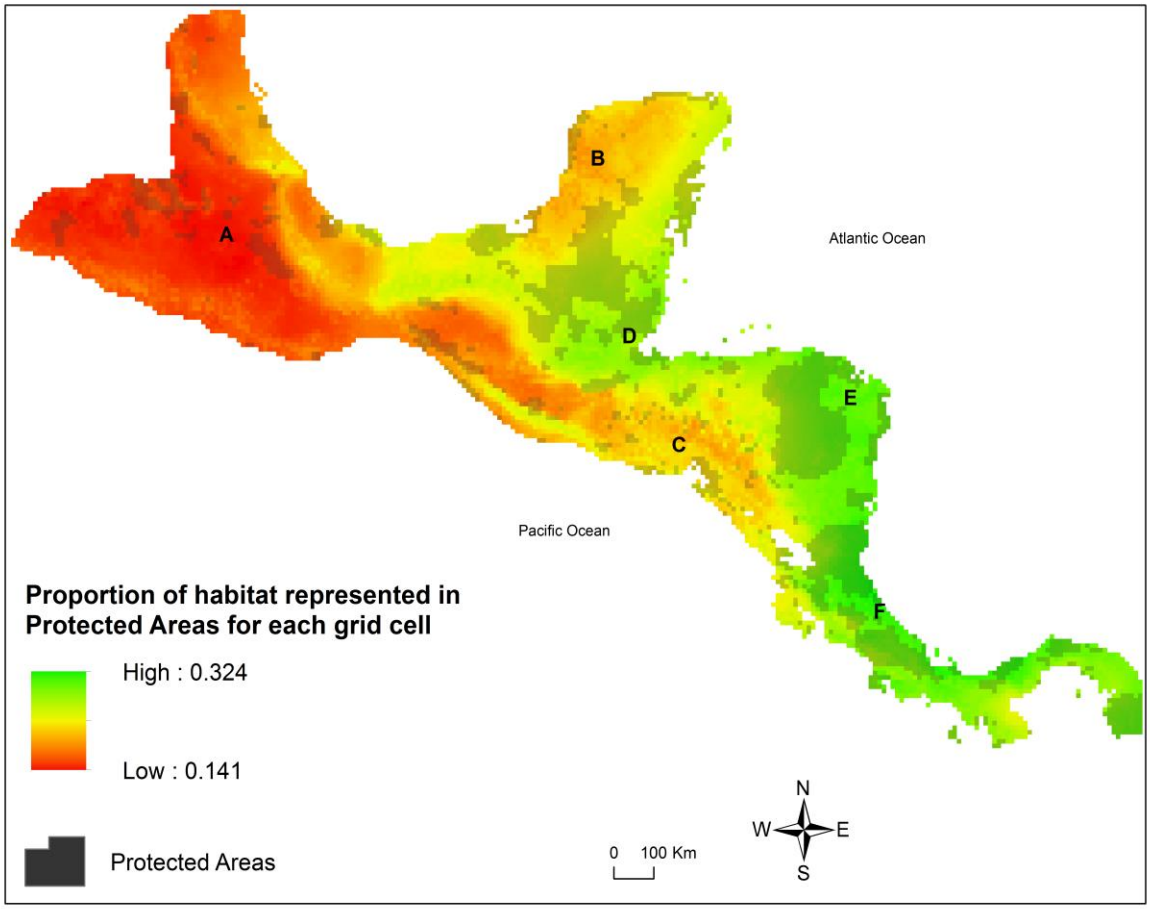


Figure 4.2 Proportion of vascular plant habitat in protected areas across Mesoamerica for the eleven predictor model, before habitat loss. Green indicates that a high proportion of areas predicted to be similar to that site are within reserves. Red indicates poor representation of similar areas within reserves.

4.3.2 Pattern of protected areas representativeness reveals geographic, altitudinal and climatic division.

The mean percentage of original habitat represented in protected areas across all ecoregions was estimated at 22.4%, with a maximum value of 30.3% for Isthmian Atlantic Moist Forests (**Figure 4.3**) and a minimum value of 14.7% for Balsas Dry Forests (**Figure 4.3**). These figures refer to protection of the assumed original habitat throughout the region. Protection coverage was also calculated for ecoregions after habitat loss up to the year 2000. When habitat loss was taken into account, the proportion of habitat protected increased almost two-fold. The proportion of protected remaining habitat varied across all ecoregions with a mean of 42.4%, a minimum of 24.8% and a maximum of 59.1% (**Appendix III, Figures 1 & 2**). However, the pattern of least and most protected ecoregions in both cases remained the same.

A total of 19 ecoregions had above average habitat representation and 16 ecoregions had below average habitat representation. **Figure 4.3** shows a clear geographic division in the proportion of habitat represented across all ecoregions. Most of the ecoregions which were better represented lie along the Atlantic coast and

in Costa Rica and Panama to the south. In contrast, most of the ecoregions which were less represented lie to the west and along the Pacific coast, with the exception of Yucatan Dry Forests in the north.

Comparing **Figure 4.3** and **Figure 4.4**, shows an altitudinal division between the least and most protected ecoregions. Ecoregions at high elevation tended to have near average or below average protection, except for Talamancan Montane Forests, which is at high elevation but had above average protection. In contrast, most of the ecoregions in lowland areas had more than average protection, except for Yucatan Dry Forests.

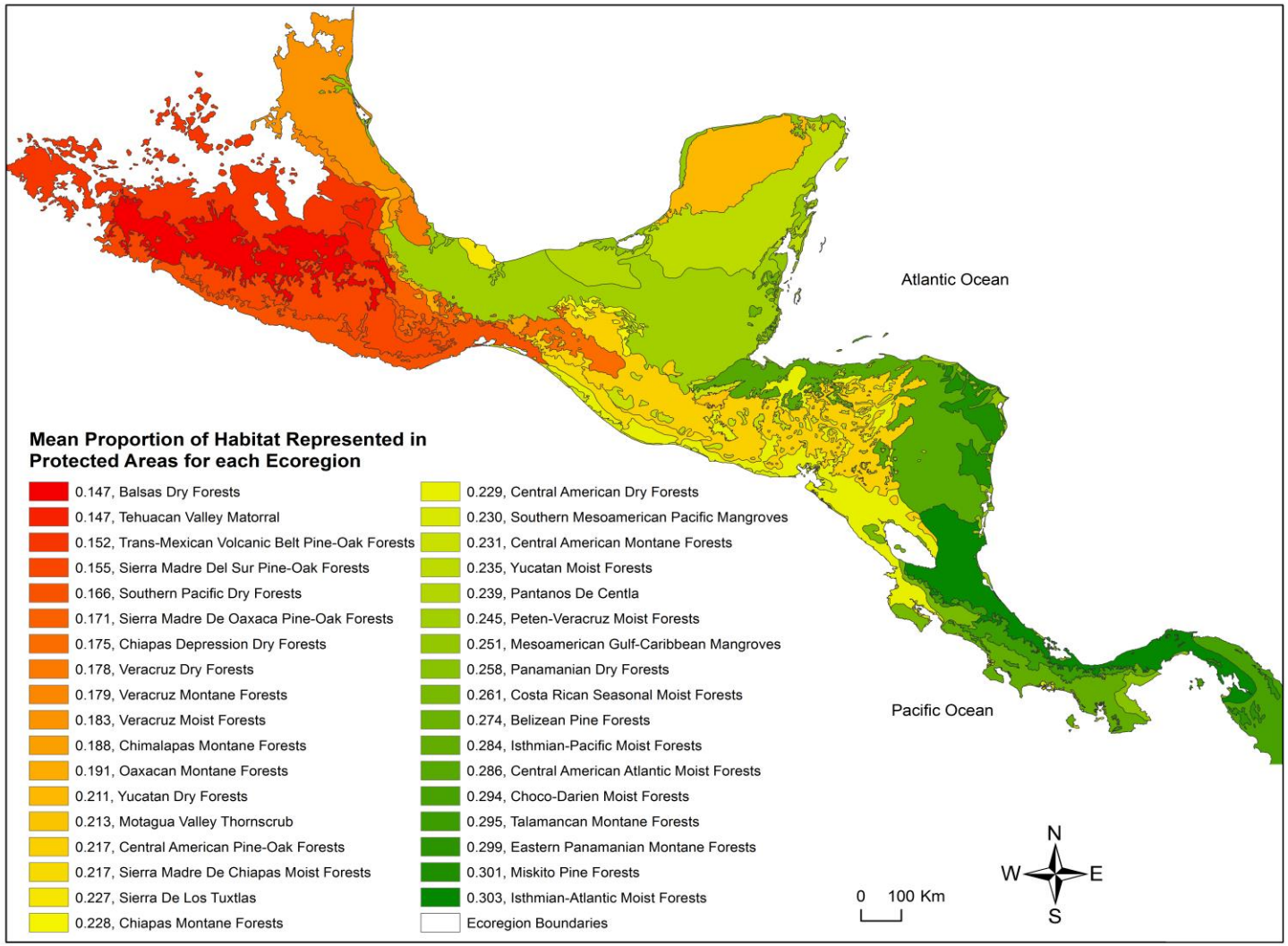


Figure 4.3 Mean proportion of original habitat represented in Protected Areas for each ecoregion. Ecoregions are sorted in ascending order of habitat representation.

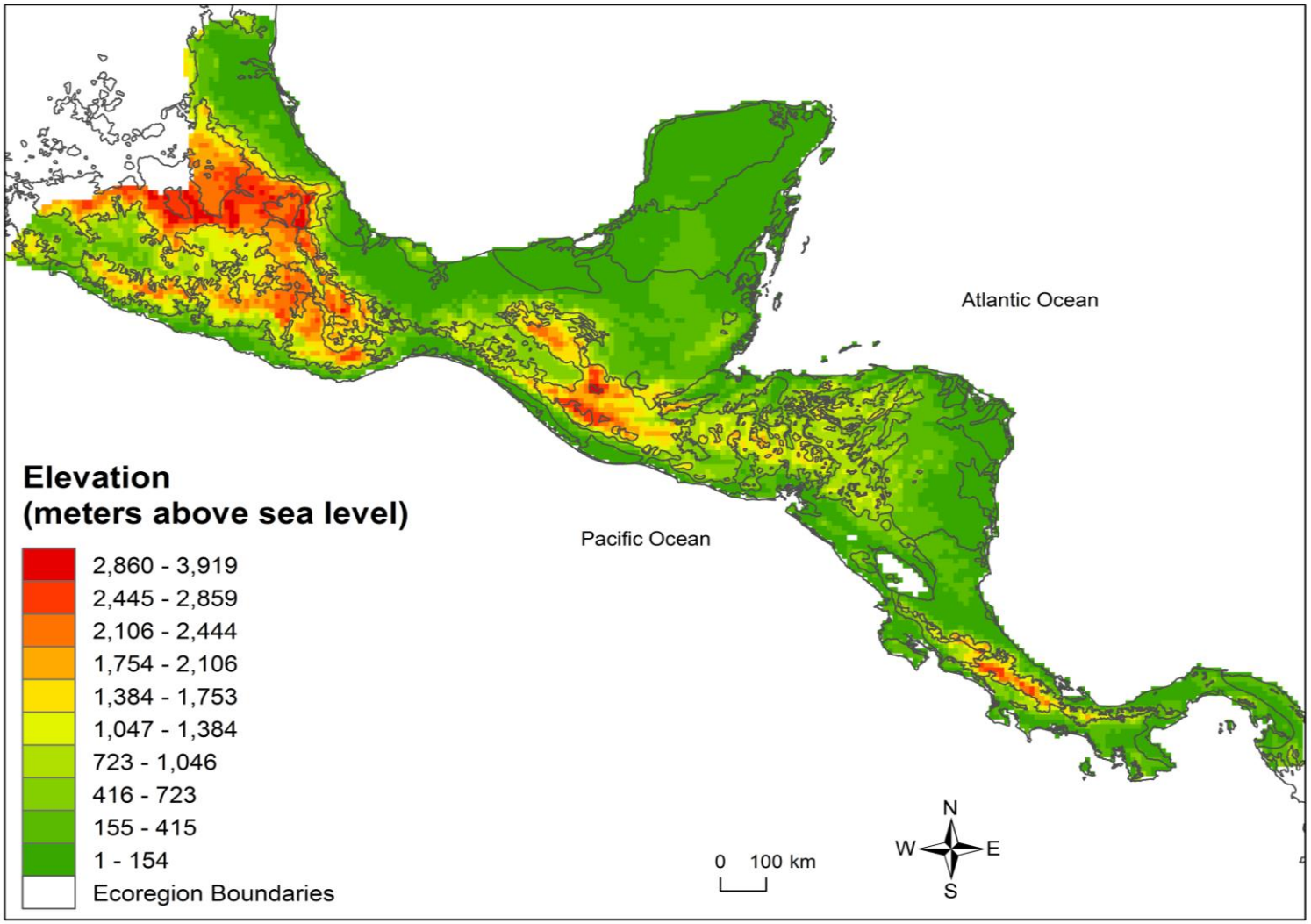


Figure 4.4 Digital Elevation Model for the Mesoamerican region.

All four Pine-Oak forest ecoregions had less than average habitat representation in protected areas (**Figure 4.5**). Two of these ecoregions were among those with the lowest proportion of habitat represented: Sierra Madre Del Sur Pine-Oak Forests (15.5%) and Trans-Mexican Volcanic Belt Pine-Oak Forests (15.2%). There are seven montane forest ecoregions in the study area. Four of these had above average representation in reserves (**Figure 4.5**): Eastern Panamanian Montane Forests (29.9%), Talamancan Montane Forests (29.5%), Central American Montane Forests (23.1%) and Chiapas Montane Forests (22.8%). The remaining three had below average representation: Oaxacan Montane Forests (19.1%), Chimalapas Montane Forests (18.8%) and Veracruz Montane Forests (17.9%). Again, this pattern showed a geographic divide as these ecoregions are located in Mexico, to the west of the study area. However, it should be noted that Central American Montane Forests and Chiapas Montane Forests have only marginally above average protection values.

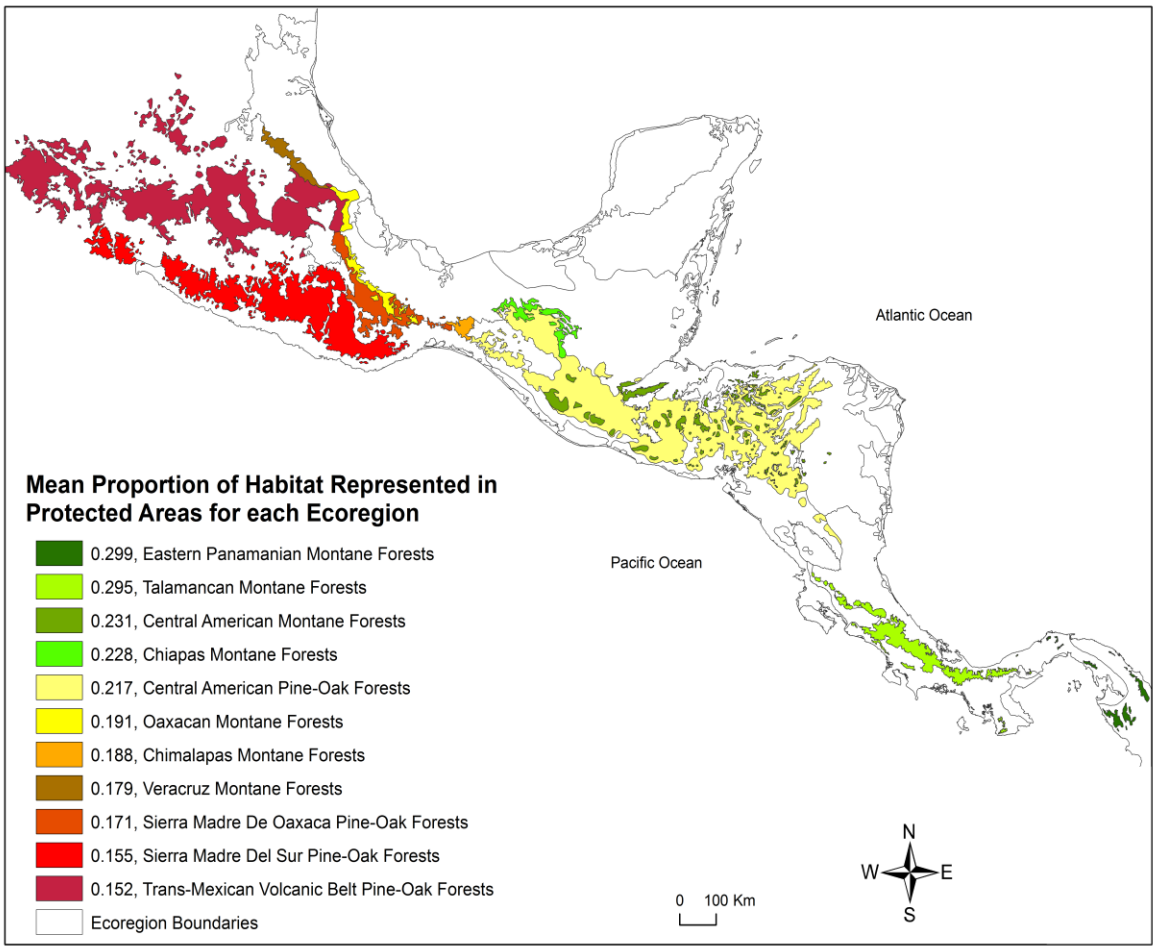


Figure 4.5 All pine-oak forest ecoregions are among those that are less represented in reserves. Some montane forest ecoregions have above average representation in reserves. Ecoregions with above average protection are shown in green shades.

There was also a clear climatic division along a precipitation gradient between moist forest ecoregions and dry forest ecoregions. Moist forest ecoregions were generally better represented in protected areas than dry forest ecoregions (**Figure 4.6**). Seven of the nine moist forest ecoregions had above average (greater than 22.4%) habitat represented. Only two moist forest ecoregions had below average habitat representation: Sierra Madre de Chiapas Moist Forests (21.7%) and Veracruz Moist Forests (18.3%). In contrast, five out of seven dry forest ecoregions had less than average representation in protected areas. Only two dry forest ecoregions had above average representation: Panamanian Dry Forests (25.8%) and Central American Dry Forests (22.9%).

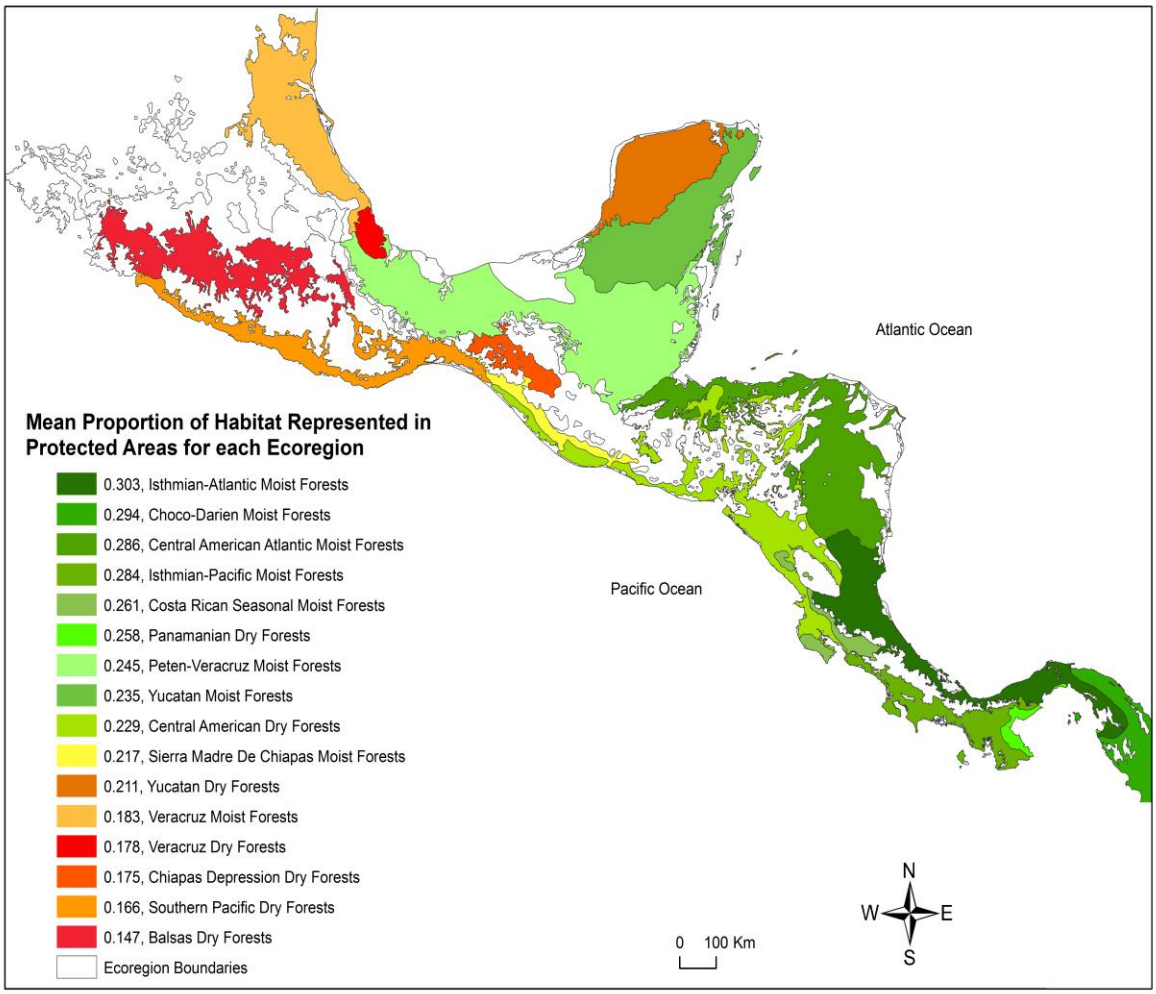


Figure 4.6 Moist forest ecoregions are better represented in protected areas than dry forest ecoregions. Ecoregions with above average protection are shown in green shades.

4.3.3 Ecoregions with highest dissimilarity are less represented in protected areas.

Mean dissimilarity across all ecoregions was 0.049 and the mean proportion of habitat in reserves across ecoregions was 0.224 (or 22.4% as stated above). Using these thresholds, results showed that 18 ecoregions had dissimilarity below the mean (**Figure 4.7**). Of these, eleven ecoregions had above average habitat represented in reserves. The remaining seven ecoregions had below average habitat representation (**Figure 4.7**).

A total of 17 ecoregions had dissimilarity equal to or above the mean of 0.049 (**Figure 4.7**). Of these, nine ecoregions had less than average habitat represented in protected areas (**Figure 4.8**). These are all located to the west of the study region. Within this set were included the top three ecoregions with the highest dissimilarities and therefore the most compositionally different from others: Tehuacan Valley Matorral, Tran-Mexican Volcanic Pine-Oak Forests, and Balsas Dry Forests. The remaining eight ecoregions had above average habitat representation (**Figure 4.8**). These lie to the middle and southeast of the study area. This set included two southern montane ecoregions: Talamancan Montane Forests and Eastern Panamanian Montane Forests.

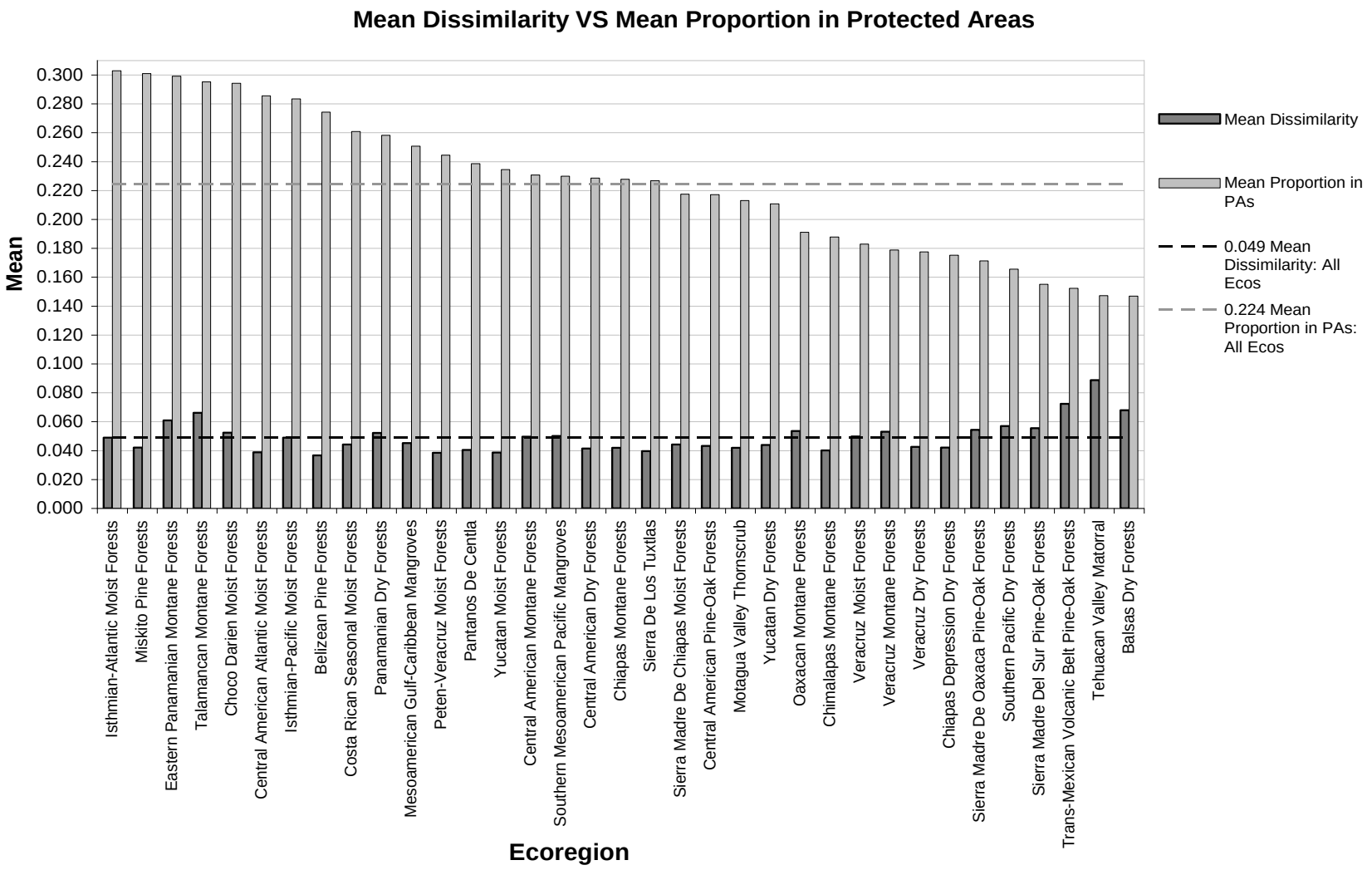


Figure 4.7 Mean dissimilarity and mean proportion of habitat represented in protected areas for each ecoregion. Ecoregions are sorted in descending order according to proportion of represented habitat. Means are calculated across all grid cells within each ecoregion. Dissimilarity values are calculated as the reciprocal of EHA (1/EHA), and the resulting value multiplied by 100, in order to fit them on the same scale.

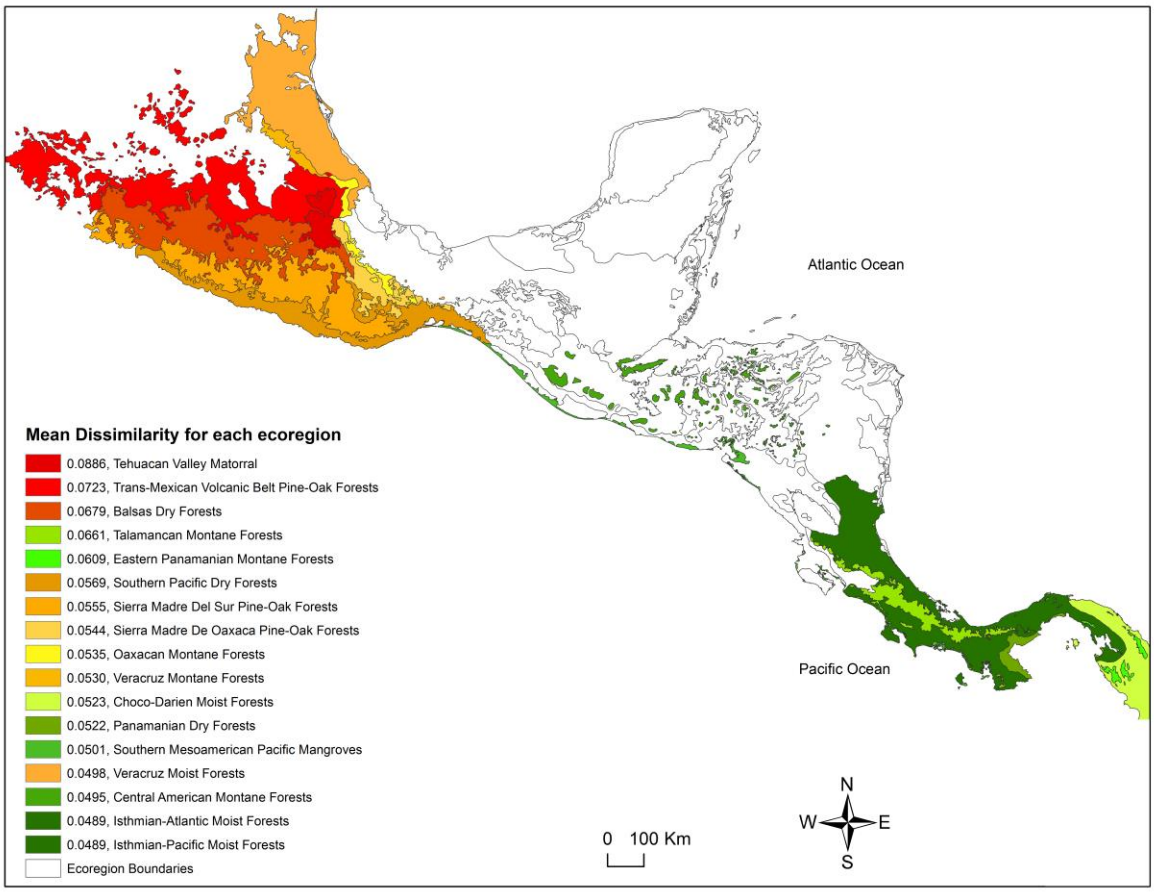


Figure 4.8 Mean dissimilarity per ecoregion. Ecoregions are colour-coded according to proportion of habitat protected in protected areas. Ecoregions with below average habitat protection are coloured in red/yellow shades. Ecoregions with above average habitat protection are indicated in green shades.

4.4 DISCUSSION

4.4.1 Representativeness of protected area system

The current regional protected area system in Mesoamerica seems to provide inadequate coverage for this known hotspot of plant diversity. This analysis estimated that a mean of 77.6% of original habitat across all ecoregions within Mesoamerica is unprotected. The proportion of unprotected original habitat varied across ecoregions between a minimum of 69.7% and a maximum of 85.3%. Using the SAR to extrapolate to species indicated that 32.3% of tree species in the assumed original species pool are unprotected by the existing protected area system. This is within the same order of magnitude as the results of a global analysis of plants and invertebrates that utilized the same GDM approach, which estimated that 43% of those species were not represented in existing protected areas (Ferrier et al. 2004).

Other global analyses have utilized different approaches to estimate the extent to which different habitats or species are represented within the global protected area network. One global analysis on vertebrates

(mammals, birds, amphibians) that used species distribution maps and Red List assessments of conservation status, identified 12-13% of species as "gap species" or unprotected species within the protected area network (Rodrigues et al. 2004a; Rodrigues et al. 2004b). The "hotspots" scheme (Myers et al. 2000), which uses endemism and degree of threat as criteria for identifying conservation priorities, indicates that 40.1% of remaining primary vegetation in the Mesoamerican hotspot is unprotected. In this study, no indication was given of the baseline year used for this measurement; it might be assumed that it considered primary vegetation up to the year 2000. The current value of 32.3% unprotected species therefore fell within the range of values reported from other studies, from 13% to 43%. However, these previous studies were conducted at a global scale whereas the current analysis was performed at a regional scale. In addition, this study focused on plants only and referred to habitat protected for those species that originally occurred in the region, or what is termed "effective habitat area". Estimates therefore did not only refer to habitat protected in terms of area but included compositional similarity, and so referred to effective habitat area protected for all species assumed to originally occur in the region. This distinction could lead to the higher estimates of unprotected habitat produced by this analysis as compared to some other estimates.

Disparities are also evident when comparing the current results with other studies that estimated protection at the ecoregion level (Soutullo et al. 2007; Jenkins and Joppa 2009). Such studies showed substantially different protection estimates than those obtained in this analysis (**Table 4.1**). However, those studies only measured how much of the area of each ecoregion was covered by protected areas. In contrast, this analysis was conducted at a finer scale (100km^2 grid cells) and was based on similarity between cells based on actual species data. The results were summarized to the ecoregion level only to facilitate comparisons. Inclusion of similarity to other cells means that estimates of protection might be relatively high because the estimate reflected the fact that the habitat in the cell of interest may also have been represented in other cells that were covered by the protected area system.

Table 4.1 Percent (%) of each ecoregion protected. Note that this table does not include all ecoregions but only those for which percentages could be assigned based on the figures in the relevant articles. Both the cited studies assessed protected area coverage of WWF terrestrial ecoregions according to Olson et al. (2001).

Ecoregion	Soutullo et al. (2007)	Jenkins & Joppa (2009)	This analysis
Veracruz Dry Forests	0	< 1	17.8
Balsas Dry Forests		1-5	14.7
Trans-Mexican Volcanic Belt Pine-Oak Forests			15.2
Sierra Madre del Sur Pine-Oak Forests			15.5
Southern Pacific Dry Forests			16.6
Sierra Madre de Oaxaca Pine-Oak Forests			17.1
Chiapas Depression Dry Forests			17.5
Veracruz Montane Forests			17.9
Oaxacan Montane Forests			19.1
Yucatan Dry Forests			21.1
Miskito Pine Forests	< 5		30.1
Central American Pine-Oak Forests	< 5	5-10	21.7
Central American Dry Forests			22.9
Central American Montane Forests			23.1

In comparing the current results to previous studies, it is important to remember that the compositional similarity predicted by the GDM model is assumed to relate to the original "pristine/primary" condition or before any recent, widespread land use change has occurred. This analysis estimates how much of the habitat within any grid cell is similar (as measured by compositional similarity) to habitat within grid cells that are protected. Therefore, the representativeness of the protected areas network refers to the proportion of original habitat represented in the current network (2010 version of the WDPA). When habitat loss up to the year 2000 is taken into account, the proportion of habitat protected increases almost two-fold. This increase in protection of remaining habitat results from the proportion of total habitat having decreased as a result of habitat loss. There was less habitat remaining in the year 2000 therefore the proportion of area protected will automatically increase simply because the overall area of available habitat has decreased. However, the overall pattern of ecoregions that are least and most protected remained the same.

4.4.2 Patterns in Representation

This study provides evidence for the biases and gaps in the protected areas system in Mesoamerica. A clear bias is shown towards protecting ecoregions of low species turnover or compositional dissimilarity, such as the moist forest ecoregions. Ecoregions with high dissimilarity are the least protected even though these ecoregions most likely contain concentrations of narrow-ranged and endemic species. A previous study employing the same methodology as this analysis, but at the global scale, showed somewhat similar findings to my results. Ferrier et al. (2004) found that the biome with the highest proportion of unrepresented biodiversity within the global protected area system was tropical and subtropical moist broadleaf forest. This was followed by tropical and subtropical grassland, savannah and shrubland biome and finally by the tropical and subtropical dry broadleaf forest biome. This research similarly showed that dry broadleaf forest

ecoregions (Balsas Dry Forest and Southern Pacific Dry Forests) are among those that are least represented in the current protected area system within Mesoamerica. In contrast, however, this analysis showed that moist broadleaf forest ecoregions (Central American Atlantic Moist Forests, Isthmian-Atlantic Moist Forest, Isthmian-Pacific Moist Forests, and Choco Darien Moist Forests) are among the most represented.

An assessment of the coverage of the global protected area network that used data on plants identified biomes that were under-represented in the system (Brooks et al. 2004). Tropical and subtropical dry broadleaf forests had low protection (9%), followed by tropical and subtropical coniferous forests (6%). This latter biome includes three of the ecoregions that this chapter has highlighted as having high dissimilarity and below average protection: Sierra Madre de Oaxaca Pine-Oak Forests, Sierra Madre del Sur Pine-Oak Forests and Trans-Mexican Volcanic Belt Pine-Oak Forests. The tropical dry forests of Mexico had the lowest coverage with less than 2% covered by the protected areas network. The current research similarly found that at the regional scale, two ecoregions within that biome have less than average protection, namely: Balsas Dry Forest and Southern Pacific Dry Forests. Results of this analysis also coincide with previous studies in identifying tropical and subtropical moist broadleaf forests as the most protected biome (Brooks et al. 2004; Chape et al. 2005). In this regional scale analysis, moist forest ecoregions also had above average protection.

Other studies have also identified ecoregions in Mexico as priority areas for conservation. For example, Hoekstra et al. (2004) applied a Conservation Risk Index (CRI) based on the ratio of per cent area converted to per cent area protected to identify ecoregions at risk. Some dry forests ecoregions in Mexico were identified as critically endangered: Sierra Madre del Sur Pine-Oak Forests, Sierra Madre de Oaxaca Pine-Oak Forests and Chiapas Depression Dry Forests. Other ecoregions such as Central American Dry Forests were identified as endangered. These ecoregions were also identified in this current analysis as being the least protected (except for Central American Dry Forests).

At the national scale, other studies have found that Costa Rica and Panama have the highest endemic species densities (Morawetz and Raedig 2007; Joppa et al. 2013) as well as the highest amount of their land in strict protection within the region (Joppa et al. 2013). In this analysis, the ecoregions in Costa Rica and Panama were also shown to have the highest proportion of protection, but have lower compositional dissimilarity when compared to the rest of the region. In addition, the current analysis coincides with local scale studies that highlighted the montane forest ecoregions and pine-oak forest ecoregions in Mexico as areas that warrant conservation effort. For example, Ponce-Reyes et al. (2012) identified areas that are important for vertebrate endemics. These areas coincided with some of the ecoregions shown in this study that are high in dissimilarity and have only average or below average protection: Veracruz Montane Forests, Sierra Madre de Oaxaca Pine-Oak Forests, Oaxaca Montane Forests, Chiapas Montane Forests, and Sierra Madre del Sure Pine-Oak Forests (**Figure 4.8**). Another study predicted that most of these ecoregions will lose more than half of their areal extent by the year 2080 owing to habitat loss and climate change (Ponce-Reyes et al. 2013). The current analysis, therefore, concurs with other findings and serves to indicate areas that are vulnerable and where protection efforts will benefit various taxa.

In general, these results for the Mesoamerican region show that high altitude ecoregions are among the least protected. This general pattern is driven mostly by the ecoregions in Mexico, which have high dissimilarity, are at high elevation and have low protection. This includes most montane, pine-oak and dry forests in Mexico, and the montane and pine-oak forests of north-western Central America. The current results agree with other studies conducted at a national scale showing the correlation between protected status and elevation (Joppa and Pfaff 2009). For example, in Mexico, there is a negative correlation between elevation and protection. These results support this finding, since most of the high elevation ecoregions in Mexico are among the least protected. The Central American Pine-Oak Forests found in Honduras and Nicaragua also showed the same pattern of below average protection with high elevation whereas Central American Montane Forests showed only slightly above average protection. Costa Rica and Panama showed a positive correlation between elevation and protection (Joppa and Pfaff 2009). The current results confirmed this pattern, in that the Talamancan Montane Forests and Eastern Panamanian Montane Forests found in these regions are among the most protected ecoregions.

Various global studies have found a general pattern of protected areas being biased towards highland areas, far from roads and urban areas, and low agricultural suitability (Oldfield et al. 2004; Gaston et al. 2008; Joppa and Pfaff 2009; Joppa and Pfaff 2010; Joppa and Pfaff 2011). This pattern is thought to reflect a bias towards siting protected areas in locations where deforestation is less likely to occur, thus reducing the impact of the protected areas network (Joppa and Pfaff 2011). Protected areas located in areas where deforestation is more likely to occur tend to correlate with flatter lands, areas closer to roads and urban areas, and lands higher in agricultural suitability. Such protected areas have a higher impact in protecting against deforestation (Joppa and Pfaff 2010). These factors were not systematically investigated in the current assessment of protected areas since the objective of this analysis was to assess how beta diversity is represented in the protected areas network, but clearly some of the results obtained could be explained by such biases.

The results of this regional analysis suggest that, when compositional dissimilarity (beta diversity) is taken into account, the ecoregions at high elevation and having above average dissimilarity (or low EHA) have the least protection. This is in contrast to the global patterns described above. In this study, there was a total of eight montane and pine-oak forest ecoregions at high elevation and with above average dissimilarity (**Figures 4.3 & 4.8**). Of these, five had below average protection. Contrary to global findings, this suggests a bias against protecting "the best areas for endemism or preventing plant extinctions" (Joppa and Pfaff 2009), since areas containing high number of endemics are likely to show low compositional similarity (Baselga 2008).

The results of this chapter confirm the gap in the current regional protected area system: areas of high compositional distinctness that may harbour endemic and range-restricted species are not well represented. Compositional similarity is likely to be low in areas with high numbers of narrowly distributed species or

endemic species. A recent global study on vertebrates (birds, mammals, amphibians) showed that richness centres for small-ranged species have more than 80% of their area unprotected (19.2% protected) (Jenkins et al. 2013). The previous study by Rodrigues et al. (2004) found that the percent of "gap species" (species not covered by any protected area) in a region was highly correlated with endemism, regardless of the percentage of area protected. Their map of gap species mainly reflected the presence of species with narrow distributions, and highlighted mostly montane and insular regions in the tropics, as well as the dry forests of Mexico. Although the current analysis was conducted at a regional rather than a global scale, the results similarly showed that ecoregions with above average compositional dissimilarity were montane ecoregions and those at high elevation, such as the dry and pine-oak forests of Mexico. More than half (53%) of these ecoregions with high dissimilarity have below average protection. This confirms the previous studies that the protected area system does not seem to cover areas that contain high numbers of endemic and narrowly distributed species.

4.4.3 Conservation Implications

These results call attention to specific habitat types that require action if areas of high distinctiveness in species composition are to be retained within the Mesoamerican region. Montane forest ecoregions, pine-oak forest ecoregions and dry forest ecoregions are mostly shown to have high dissimilarity. Most of these ecoregions have low protection (**Figure 4.8**). Some of them, such as Chiapas Montane Forests, Chimalapas Montane Forests, Chiapas Depression Dry Forests, Central American Pine-Oak Forests, Central American Dry Forests and Yucatan Dry Forests have intermediate mean dissimilarity (only slightly below average) (**Figure 4.7**). Most of these have below average protection, except for Chiapas Montane Forests and Central American Dry Forests, which have only slightly above average protection (**Figure 4.3**). These results reinforce the findings of other studies that show montane, dry and pine-oak forests as areas of high diversity and vulnerability.

Previous studies have identified montane forests in the Chiapas Montane Forests ecoregion as highly fragmented centres of tree diversity (Cayuela et al. 2006). Montane forests can degrade to oak or pine-oak forests as a result of selective logging practices (Cayuela et al. 2006). Thus pine-oak forests could serves as refuges for seedlings of other species and, if not degraded further, could undergo succession to montane forests. Threats to montane forests include land cover change, logging, fragmentation, infrastructure development, and climate change. Miles et al. (2007) projected the major pressures to affect Veracruz Montane Forests to be logging in 2010 and fragmentation and climate change in 2050. The Chiapas Montane Forests were predicted to be mostly threatened by land cover change, fragmentation and logging in 2010 and by fragmentation and logging in 2050. An integrated assessment of pressures including fire risk, infrastructure development, agricultural conversion and climate change found that the Chiapas ecoregions (Chiapas Montane Forests, Chiapas Depression Dry Forests), Central American Montane Forests and Central American Pine-Oak Forests ecoregions are threatened by up to three forms of sever pressure (Blyth et al.

2002). That study also found that the presence of protected areas in montane regions is lower in southern Mexico than in Central America, a finding that is supported by this analysis, which showed the montane ecoregions in Mexico as being among the least protected. Other studies have also recommended the conservation of the cool, moist microclimates restricted to mountain peaks and ridges in Mexico and Central American, in order to slow species loss (Golicher et al. 2008). In addition, the effect of climate change on montane forests has already been documented in Costa Rica (Foster 2001). For example, decreases in humidity in montane forests due to a reduction in cloud immersion can be fatal for characteristic and sensitive montane forest communities such as epiphytes, which occupy narrow ecological niches (Foster 2001). Changes such as these make montane forest species particularly vulnerable since they may not be able to migrate due to a lack of locations with suitable climate.

The dry forests of Mexico and Central America are of high biodiversity value according to my results of high compositional dissimilarity. They are also highly threatened from fire risk, human encroachment owing to high population density, fragmentation and climate change. Miles et al. (2006) found dry forests in Latin America had experienced the greatest decrease in percentage area forested between 1980 and 2000. Mexican dry forests were found to have high fire risk. North and Central American dry forests had the highest population densities, high fragmentation and low percentage of protection (Miles et al. 2006). In addition, Mesoamerican dry forests have been extensively cleared as a result of their relatively fertile soils and climate favourable to human habitation (Wright and Muller-Landau 2006).

Climate change projections for the region predict large decreases in precipitation varying from 25% to 39% for eastern and southern Mexico, the Yucatan Peninsula and Central America (Karmalkar et al. 2011). With such decreases in precipitation, it is possible that new areas could become climatically suitable for dry forests (Miles et al. 2006). The timescale over which such new colonisations could occur is not known, so it might be wise to apply the precautionary principle and protect these areas. Many species in dry forests have been found to exhibit strong adaptations and fidelity to the ecological characteristics of the dry forest habitat (i.e. niche conservatism) (Pennington et al. 2000, 2009). This means that it is more probable for dry forest lineages to disperse to other distant dry forest areas, than for species from nearby biomes to develop the necessary adaptations to colonize dry forest areas (Pennington et al. 2009). These observations suggest that dry forests may represent communities that are more resilient to climate change in the short-term and adaptable in the long-term. They may serve as evolutionary refugia, being a source of immigrants over time (Pennington et al. 2009). It is therefore important that dry forests be a high priority for conservation efforts.

The patterns of compositional turnover obtained from the GDM model could help locate protected areas along spatial gradients that maximize species diversity, as has been shown in other areas, e.g. the Yukon in Canada (Wiersma and Urban 2005). The northwestern part of the study region and the Chiapas region showed a uniform pattern of cells of low EHA or high compositional dissimilarity (**Chapter 2, Figure 2.2**). In these parts of the region, it may be more effective to create large reserves that capture these areas of high species turnover. Other areas, however, showed a less uniform or patchier pattern of species turnover. In

these areas, it might be more efficient to create a network of smaller reserves. This is especially true in the case of Central American Montane and Central American Dry Forests, which are confined to fragmented pockets. Finally, the Atlantic coast of the region showed areas with little change in species composition, high EHA or low compositional dissimilarity. These areas may only require a single large reserve to capture the full set of species. In this regard, it is clear that the regional protected areas network has a strong southward bias (**Figure 4.8**), and a strong bias towards moist forest ecoregions, as these areas had the highest proportion of habitat protected. This suggests that many of the current protected areas are protecting more of the same thing, as these regions have the lowest turnover (high EHA).

Conservation priority assessments ideally should combine studies of diversity patterns with localized studies that can address other factors operating at more local scales. In the current analysis, the Oaxacan Montane Forests ecoregion was shown to be of high dissimilarity but have low representation in protected areas. This would suggest that this entire region presents an opportunity for conservation of this highly threatened ecosystem. Local factors, however, also influence which areas present opportunities for conservation. In the case of the Oaxacan Montane Forests ecoregion, there are mixed opportunities for conservation based on forest quality, threats to forest permanence, threats to forest integrity, and opportunities for conservation (Toledo-Aceves et al. 2011). The ecoregion can be divided into four subregions, each having a different prioritisation varying from high to critical. Some subregions have large areas of high quality forest and high threat but low opportunities for conservation due to land tenure conflicts, whereas others have smaller, more fragmented forest and low threat but better opportunities for conservation (Toledo-Aceves et al. 2011). Such trade-offs are inevitable aspects of conservation planning, and such studies demonstrate that while regional analyses on diversity patterns can provide a valuable starting point, other more local factors must be considered in prioritizing conservation efforts and funding.

4.4.4 Caveats

Beta diversity or compositional dissimilarity can be used to assess the representativeness of the system of protected areas. In this analysis, the concept of compositional dissimilarity, applied directly to tree species data, shows distinct patterns that coincide with some previous studies. Differences with these other studies are found, however, in the estimates of the degree of protection. At the ecoregion scale, for example, all ecoregions showed a level of protection above the 10% CBD target. This is in contradiction to other studies that have found levels of protection below 10% throughout the region. The difference is most likely due to the different method utilized in this study: the incorporation of compositional dissimilarity in deriving estimates of protection.

Unlike other studies, this analysis considers the similarity of grid cells to other grid cells in calculating an estimate of protection. If the habitat in one cell is protected and is also found protected in another cell, then they both contribute equally to the proportion of that habitat that is protected. In this case, the habitat in that cell will have a higher proportion of habitat protected than a cell whose habitat is not found in any other cell. At the ecoregion scale, this means that ecoregions whose habitat is also found in other ecoregions will have a

higher proportion of representation in protected areas. Other studies do not incorporate this concept (except see Ferrier et al. 2004) but instead treat ecoregions as distinct areas that are equally different from one another. The consideration of compositional similarity/dissimilarity (beta-diversity) therefore leads to higher estimates of protection in this analysis than those found in other studies.

It should be noted that the WDPA protected areas raster layer used includes all designations of protected areas. This study does not distinguish between protected areas in terms of management effectiveness or enforcement level. No distinction is made between strict protected areas with no extraction activities, and designations, such as, forest reserves that allow timber harvesting. All protected areas are treated equally and assumed to have the same degree of protection. In addition, the raster layer is a continuous raster of grid cells with proportions of protection that vary from 0 to 1. This would lead to higher estimates of mean protection at the ecoregion scale when compared to other studies that use only a binary distinction of 0 and 1 (not protected versus protected).

4.5 CONCLUSION

In summary, this analysis provides a regional perspective on where tree species diversity is greatest and may be at greatest risk due to insufficient protection. It identifies ecoregions where future conservation efforts should be focused. Such efforts should focus on areas containing range-restricted or endemic species. This is shown in this analysis by areas of high compositional dissimilarity. Prioritizing such areas may also prioritize locations where species will be at future risk, e.g. montane forests where species are at risk of climate change. So far, it is evident from this analysis that the level of protection in the region is inadequate with more than three-quarters of total habitat left unprotected, including many areas with the highest compositional dissimilarity. Conservation efforts have focused on areas of low compositional dissimilarity and have focused mainly on moist forests. This bias needs to be addressed if the vast plant diversity of the Mesoamerican region will be adequately represented and preserved in the future.

Specific recommendations arising from this research include:

- Increase the level of protection in the region. A mean level of protection of less than 25% across all ecoregions seems inadequate, considering that the entire region is a centre of plant diversity and is considered at high risk for land conversion, fragmentation, increased population density and climate change.
- Address the bias towards protecting more tropical moist forests by focusing efforts within those ecoregions shown to have compositional distinctness and high risk, especially montane forests, dry forests and pine-oak forests.
- Address the bias in protected areas coverage by focusing more conservation efforts within other countries in the region that contain less protected ecoregions.

4.6 REFERENCES

Allnutt, T. F., Ferrier, S., Manion, G., Powell, G. V. N., Ricketts, T. H., Fisher, B. L., Harper, G. J., Irwin, M. E., Kremen, C., Labat, J.-N., Lees, D. C., Pearce, T. a. and Rakotondrainibe, F., 2008. A method for quantifying biodiversity loss and its application to a 50-year record of deforestation across Madagascar. *Conservation Letters*, *1* (4), 173–181.

Baselga, A., 2008. Determinants of species richness, endemism and turnover in European longhorn beetles. *Ecography*, *31* , 263–271.

Blyth, S., Groombridge, B., Lysenko, I., Miles, L. and Newton, A. C., 2002. *Mountian Watch*. *Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki*. Cambridge, UK.

Brooks, T. M., Bakarr, M. I., Boucher, T., Da Fonseca, G. a. B., Hilton-Taylor, C., Hoekstra, J. M., Moritz, T., Olivieri, S., Parrish, J., Pressey, R. L., Rodrigues, A. S. L., Sechrest, W., Stattersfield, A., Strahm, W. and Stuart, S. N., 2004. Coverage Provided by the Global Protected-Area System: Is It Enough? *BioScience*, *54* (12), 1081-1091.

Cayuela, L., Benayas, J. M. R. and Echeverría, C., 2006. Clearance and fragmentation of tropical montane forests in the Highlands of Chiapas, Mexico (1975–2000). *Forest Ecology and Management*, *226* (1-3), 208–218.

Chape, S., Harrison, J., Spalding, M. and Lysenko, I., 2005. Measuring the extent and effectiveness of protected areas as an indicator for meeting global biodiversity targets. *Philosophical Transactions of the Royal Society of Biological Sciences*, *360* (1454), 443–455.

Ferrier, S., Harwood, T., Williams, K. J. and Dunlop, M., 2012. *Using Generalised Dissimilarity Modelling to assess potential impacts of climate change on biodiversity composition in Australia , and on the representativeness of the National Reserve System Climate Adaptation Flagship Working Paper # 13E*. Canberra.

Ferrier, S., Powell, G. V. N., Richardson, K. S., Manion, G., M, J., Allnutt, T. F., Cameron, S. E., Mantle, K., Burgess, N. D., Faith, D. P., Lamoreux, J. F., Kier, G., Hijmans, R. J., Funk, V. A., Cassis, G. A., Fisher, L., Flemons, P., Lees, D., Lovett, J. C., Rompaey, R. S. A. R. Van, Overton, J. M., Allnutt, F., Fisher, B. L., Lovett, J. O. N. C. and Rompaey, R. S. A. R. V. A. N., 2004. Mapping More of Terrestrial Global Biodiversity for Conservation Assessment. *BioScience*, *54* (12), 1101–1109.

Foster, P., 2001. The potential negative impacts of global climate change on tropical montane cloud forests. *Earth-Science Reviews*, *55* , 73–106.

Gaston, K. J., Jackson, S. F., Cantú-Salazar, L. and Cruz-Piñón, G., 2008. The Ecological Performance of Protected Areas. *Annual Review of Ecology, Evolution, and Systematics*, *39* (1), 93–113.

Golicher, D. J., Cayuela, L., Alkemade, J. R. M., González-Espinosa, M. and Ramírez-Marcial, N., 2008. Applying climatically associated species pools to the modelling of compositional change in tropical montane forests. *Global Ecology and Biogeography*, *17* (2), 262–273.

Hoekstra, J. M., Boucher, T. M., Ricketts, T. H. and Roberts, C., 2004. Confronting a biome crisis: global disparities of habitat loss and protection. *Ecology Letters*, *8* (1), 23–29.

Jenkins, C. N. and Joppa, L., 2009. Expansion of the global terrestrial protected area system. *Biological Conservation*, *142* (10), 2166–2174.

Jenkins, C. N., Pimm, S. L. and Joppa, L. N., 2013. Global patterns of terrestrial vertebrate diversity and conservation. *Proceedings of the National Academy of Sciences of the United States of America*, *110* (28), e2602–2610.

Joppa, L. N. and Pfaff, A., 2009. High and Far: Biases in the Location of Protected Areas. *PLoS ONE*, *4* (12), e8273.

Joppa, L. N. and Pfaff, A., 2010. Reassessing the forest impacts of protection: the challenge of nonrandom location and a corrective method. *Annals of the New York Academy of Sciences*, *1185* , 135–149.

Joppa, L. N. and Pfaff, A., 2011. Global protected area impacts. *Proceedings of the Royal Society of Biological Sciences*, *278* (1712), 1633–1638.

Joppa, L. N., Visconti, P., Jenkins, C. N. and Pimm, S. L., 2013. Achieving the convention on biological diversity's goals for plant conservation. *Science*, *341* (6150), 1100–1103.

Karmalkar, A. V., Bradley, R. S. and Diaz, H. F., 2011. Climate change in Central America and Mexico: regional climate model validation and climate change projections. *Climate Dynamics*, *37* (3-4), 605–629.

Koleff, P., Gaston, K. J. and Lennon, J. J., 2003. Measuring beta diversity for presence-absence data. *Journal of Animal Ecology*, *72* (3), 367–382.

Margules, C. R. and Pressey, R. L., 2000. Systematic conservation planning. *Nature*, *405* (6783), 243–253.

Miles, L., Newton, A. C., Alvarez-Aquino, C., Armesto, J. J., Del Castillo, R. F., Cayuela, L., Echeverría, C., González-Espinosa, M., Lara, A., Lopez-Barrera, F., Manson, R. H., Montoya-Gomez, G., Muniz-Castro, M. A., Nunez-Avila, M. C., Pedraza, R. A., Rey Benayas, J. M., Rovere, A. E., Ruger, N., Smith-Ramirez, C., Souto, C. and Williams-Linera, G., 2007. Future scenarios for tropical montane and south temperate forest biodiversity in Latin America. *In*: Newton, A. C., ed. *Biodiversity loss and conservation in fragmented forest landscapes: the forests of montane Mexico and temperate South America*. Wallingford, United Kingdom: CAB International, 370–398.

Miles, L., Newton, A. C., DeFries, R. S., Ravilious, C., May, I., Blyth, S., Kapos, V. and Gordon, J. E., 2006. A global overview of the conservation status of tropical dry forests. *Journal of Biogeography*, *33* (3), 491–505.

Morawetz, W. and Raedig, C., 2007. Angiosperm Biodiversity , Endemism and Conservation in the Neotropics. *Taxon*, *56* (4), 1245–1254.

Myers, N., Mittermeier, R. a, Mittermeier, C. G., da Fonseca, G. a and Kent, J., 2000. Biodiversity hotspots for conservation priorities. *Nature*, *403* (6772), 853–858.

Oldfield, T. E. ., Smith, R. J., Harrop, S. R. and Leader-Williams, N., 2004. A gap analysis of terrestrial protected areas in England and its implications for conservation policy. *Biological Conservation*, *120* (3), 303–309.

Pennington, R. T., Lavin, M. and Oliveira-Filho, A., 2009. Woody Plant Diversity, Evolution, and Ecology in the Tropics: Perspectives from Seasonally Dry Tropical Forests. *Annual Review of Ecology, Evolution, and Systematics*, *40* (1), 437–457.

Pennington, R. T., Prado, D. E., Pendry, C. A. and Botanic, R., 2000. Neotropical seasonally dry forests and Quaternary vegetation changes. *Journal of Biogeography*, *27* , 261–273.

Ponce-Reyes, R., Nicholson, E., Baxter, P. W. J., Fuller, R. a. and Possingham, H., 2013. Extinction risk in cloud forest fragments under climate change and habitat loss. *Diversity and Distributions*, *19* (5-6), 518–529.

Ponce-Reyes, R., Reynoso-Rosales, V.-H., Watson, J. E. M., VanDerWal, J., Fuller, R. a., Pressey, R. L. and Possingham, H. P., 2012. Vulnerability of cloud forest reserves in Mexico to climate change. *Nature Climate Change*, *2* (6), 448–452.

Pressey, R. L., Cabeza, M., Watts, M. E., Cowling, R. M. and Wilson, K. A., 2007. Conservation planning in a changing world. *Trends in Ecology & Evolution*, *22* (11), 583–592.

Prober, S. M., Hilbert, D. W., Ferrier, S., Dunlop, M. and Gobbett, D., 2012. Combining community-level spatial modelling and expert knowledge to inform climate adaptation in temperate grassy eucalypt woodlands and related grasslands. *Biodiversity and Conservation*, *21* (7), 1627–1650.

Rodrigues, A., Andelman, S., Bakarr, M., Boitani, L., Brooks, T. M., Cowling, R. M., Fishpool, L. D. C., da Fonseca, G. A. B., Gaston, K. J., Hoffmann, M., Long, J. S., Marquet, P. A., Pilgrim, J. D., Pressey, R. L., Schipper, J., Sechrest, W., Stuart, S. N., Underhill, L. G., Waller, R. W., Watts, M. E. J. and Yan, X., 2004a. Effectiveness of the global protected area network in representing species diversity. *Nature*, *428* (April), 640–643.

Rodrigues, A. S. L., Akçakaya, H. R., Andelman, S. J., Bakarr, M. I., Boitani, L., Brooks, T. M., Chanson, J. S., Fishpool, L. D. C., Da Fonseca, G. a. B., Gaston, K. J., Hoffmann, M., Marquet, P. a., Pilgrim, J. D., Pressey, R. L., Schipper, J., Sechrest, W., Stuart, S. N., Underhill, L. G., Waller, R. W., Watts, M. E. J. and Yan, X., 2004b. Global Gap Analysis: Priority Regions for Expanding the Global Protected-Area Network. *BioScience*, *54* (12), 1092-1100.

Soutullo, A., De Castro, M. and Urios, V., 2007. Linking political and scientifically derived targets for global biodiversity conservation: implications for the expansion of the global network of protected areas. *Diversity and Distributions*, *14* (4), 604–613.

Toledo-Aceves, T., Meave, J. a, González-Espinosa, M. and Ramírez-Marcial, N., 2011. Tropical montane cloud forests: current threats and opportunities for their conservation and sustainable management in Mexico. *Journal of Environmental Management*, *92* (3), 974–981.

Tuomisto, H. and Ruokolainen, K., 2006. Analyzing or explaining beta diversity? Understanding the targets of different methods of analysis. *Ecology*, *87* (11), 2697–2708.

Wescott, G., 2006. The long and winding road: The development of a comprehensive, adequate and representative system of highly protected marine protected areas in Victoria, Australia. *Ocean & Coastal Management*, *49* (12), 905–922.

Wiersma, Y. F. and Urban, D. L., 2005. Beta Diversity and Nature Reserve System Design in the Yukon, Canada. *Conservation Biology*, *19* (4), 1262–1272.

Williams, K. J., Ferrier, S., Rosauer, D., Yeates, D., Manion, G., Harwood, T., Stein, J., Faith, D. P., Laity, T. and Whalen, A., 2010. *Harnessing Continent-Wide Biodiversity Datasets for Prioritising National Conservation Investment*. *Communities*. Canberra, Australia.

Wilson, K. a, Underwood, E. C., Morrison, S. a, Klausmeyer, K. R., Murdoch, W. W., Reyers, B., Wardell-Johnson, G., Marquet, P. a, Rundel, P. W., McBride, M. F., Pressey, R. L., Bode, M., Hoekstra, J. M., Andelman, S., Looker, M., Rondinini, C., Kareiva, P., Shaw, M. R. and Possingham, H. P., 2007. Conserving biodiversity efficiently: what to do, where, and when. *PLoS Biology*, *5* (9), e223.

Wright, S. and Muller-Landau, H., 2006. The Future of Tropical Forest Species. *Biotropica*, *38* (3), 287–301.

5 CHAPTER 5

Assessing similarity of ecoregion units in Mesoamerica

5.1 INTRODUCTION

Prioritizing areas for conservation is an important task in conserving natural communities and the ecosystem services they provide. Recent years have been an exciting time in ecology as the fields of biogeography and conservation come together to tackle issues relevant to the persistence of the planet's biodiversity (Whittaker et al. 2005). The increasing availability of data (Hill et al. 2010) provides a crucial opportunity to study the distributional patterns of species and apply emerging insights to conservation planning efforts. Former planning frameworks can be assessed in light of new data to ensure their representativeness in depicting diversity patterns.

One planning approach which is used extensively in conservation priority-setting is the ecoregion (Giam et al. 2011). Ecoregions were originally subdivided from larger biogeographic realms to represent relatively homogenous units of distinct species assemblages (Mcdonald et al. 2005; Olson and Dinerstein 1998). However these ecoregions were defined on the basis literature reviews and expert consultation (Miles et al. 2006). The availability of data and advances in the study of diversity patterns offer new ways to apply data-based frameworks to the delineation of biogeographic regions (Barthlott et al. 2007; Kreft and Jetz 2010) and to evaluate previous strategies such as the ecoregion approach (Kier et al. 2005).

The recent focus on applying species turnover (beta-diversity) as a useful method of measuring diversity and differentiating between areas of interest (Anderson et al. 2011; Baselga 2010; Bridgewater et al. 2004; Chase 2010; Condit et al. 2002; Kraft et al. 2011) can be applied at the ecoregion scale to determine whether these units capture species turnover patterns (McDonald et al. 2005). Here I evaluate whether the ecoregions classification system captures similarity in species composition (species turnover or beta diversity) among ecoregions by addressing the following questions:

1. How does species compositional similarity vary among ecoregions?

2. What is the contribution of spatial distance, mean annual temperature and mean annual precipitation as potential drivers of variation in species turnover among ecoregions?

5.2 METHODS

5.2.1 Species Data

For this analysis, a species list for the region was collated based on species identified within inventory plots across the region as part of the Tree Biodiversity Network (BIOTREE-NET) initiative (Cayuela et al. 2012a). There were a total of 2001 inventory plots across the region. Within these plots, mainly angiosperms were reported as woody trees and shrubs (height > 4m; most 2.5-10cm dbh). The ecoregions with the least number of plots, having less than 10 plots each were: Yucatan Dry Forests, Southern Mesoamerican Pacific Pacific Mangrove Forests, Belizean Pine Forests, Veracruz Moist Forests, Sierra Madre del Sure Pine-Oak Forests, Eastern Panamanian Montane Forests, Miskito Pine Forests, Sierra de lot Tuxtlas. The ecoregion with the highest number of plots was Central American Pine-Oak Forests followed by Isthmian Atlantic Moist Forests. No plots were registered in Guatemala since no information was obtained from that country.

A list was created for all angiosperm species recorded within all inventory plots. This species list was then used to download geo-referenced specimen data from the Global Biodiversity Information Facility (GBIF) for the entire study region (**Figure 5.1**). The dataset underwent taxonomic standardization using the R package TAXONSTAND Version 1.2, which standardizes plant names using The Plant List (http://www.theplantlist.org) to retrieve information about each species' taxonomic status (Cayuela et al. 2012b). The Plant List is a collaborative effort among botanic gardens and draws on information about plant nomenclature from various sources. For angiosperms, the Plant List assigns accepted genera to the families recognised by the Angiosperm Phylogeny Group according to the 2009 updated family and subclass level classifications (Chase and Reveal 2009). Standardization of the dataset showed that 50.7% of species names in the dataset were accepted names, 11.2% of names were synonyms, 1.7% had unresolved names, and 36.4% had names for which a determination could not be made on accepted, synonym or unresolved (Cayuela et al. 2012b).

It is important to note that an absence of inventory plots in an area (e.g. in Guatemala) does not imply that no species data was available for that area in this analysis. This is because the plots were only used to collate a species list for the region that was then used to download specimen record data from across the entire study region. The final dataset used for this analysis consisted of over 478,774 geo-referenced specimen records from GBIF, representing 2,561 species of mostly woody trees and shrubs.

The data were stratified by ecoregions to obtain a species list for all ecoregions within the study area. This was achieved by overlaying the distribution data on a map of ecoregions obtained from the World Wildlife Fund website (http://www.worldwildlife.org/science/data/item6373.html) developed by Olson et al. (2001). The ecoregions selected for inclusion in this study were all those between 22 and 1.5° latitude. The geographic coordinate for each ecoregion was taken to be the centre of each ecoregion polygon. I included ecoregions which went beyond the boundaries of the study region proper in order to be able to observe similarities with more northern or southern ecoregions. These included 20 ecoregions extending into Mexico and one ecoregion extending into South America (the Choco-Darien Moist forest). However, I excluded ecoregions with less than 90 records (Tehuacan Valley Matorral) and Lake

Nicaragua. A total of 34 ecoregions were used in the analysis.

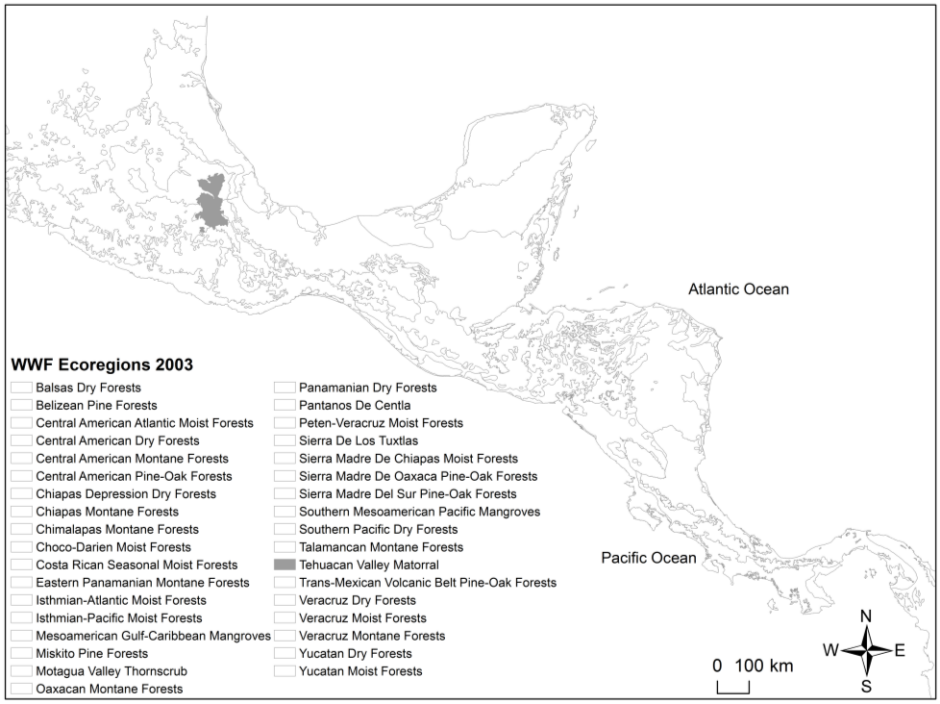


Figure 5.1 Map of study region showing boundaries of ecoregions included in the study. Tehuacan Valley Matorral (gray) and bodies of water were excluded from the analysis.

Table 5.1 summarizes the final dataset (Appendix IV summarizes the data per ecoregion). The ecoregion with the most records was Isthmian-Atlantic Moist forest with 76,503 records, while the ecoregion with the least records was Montagua Valley Thornscrub with 704 records. Some of the most widespread families, that is, those families which occurred in all 34 ecoregions, included Malphigaceae, Cannabaceae, Araliaceae, Rubiaceae and Leguminosae. As **Table 5.1** shows, 20% of species occurred in at least half (17) of all ecoregions. More than half (58%) of all species occurred in less than 10 ecoregions.

Table 5.1 Summary of data used in the study

Sampling unit	Total number	No. of taxa occupying>16 ecoregions	No. of taxa occupying <10 ecoregions
Record	478,774	------	-----
Species	2,561	534 (20%)	1556 (58%)
Genus	831	368 (44%)	286 (34%)
Family	153	115 (75%)	15 (10%)

5.2.2 Climate data

Climate data were obtained from the WorldClim Global Climate data website (www.worldclim.org) (Hijmans et al. 2005). Annual precipitation (in mm) and mean annual temperature (in ^{0}C) were derived

for each ecoregion. Mean annual temperature was derived by averaging the minimum and maximum temperatures for each month (Hijmans et al. 2005).

5.2.3 Data Analysis

5.2.3.1 Assessing Variation in Species Turnover

5.2.3.1.1 Cluster Analysis: Ecoregions

In order to assess the similarity in species composition among ecoregions, hierarchical cluster analysis was performed based on the species data for each ecoregion using the BIODIVERSE software package v. 0.16beta (http://purl.org/biodiverse) (Laffan et al. 2010). In this case, the ecoregions were considered as the sampling sites, which were used to create a site by species binary matrix (1 represents presence; 0 represents absence) (**Figure 5.2**). Based on the site by species matrix, the amount of species turnover between each pair of sites was estimated using an appropriate distance metric (**Table 1.1**). In this case, I used the Simpson dissimilarity index, B_{sim}, as it is independent of richness differences between sites (Koleff et al. 2003). The resulting distance matrix provides B_{sim} dissimilarity values for all pair-wise combination of sites. Cluster analysis was then applied to the distance/dissimilarity matrix in order to form groups of sites with similar species assemblages (Kreft and Jetz 2010). The clustering procedure used was agglomerative clustering based on the unweighted pair-group method using arithmetic averages (UPGMA or average linkage method), which has been shown to perform better in other biogeographic analyses (Kreft and Jetz 2010). The results of the cluster analysis were displayed as dendrograms and maps.

Site by species matrix

	Species A	Species B	Species B
Site a	1	0	1
Site b	0	1	1
Site c	1	0	0

Distance matrix

	Site a	Site b	Site c
Site a	0	D*ab*	D*ac*
Site b	D*ab*	0	D*bc*
Site c	D*ac*	D*bc*	0

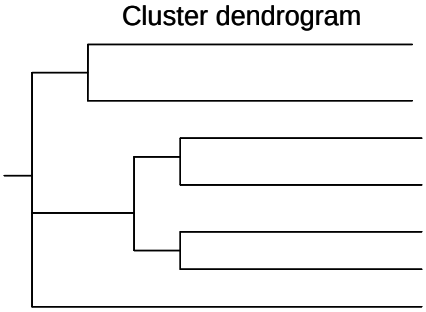


Figure 5.2 Cluster analysis process. The presence/absence matrix is used to create a distance/dissimilarity matrix of pair-wise B_{sim} dissimilarities (D) between all combinations of sites. Only the upper or lower diagonal values are used as they are both the same. Cluster analysis is then conducted based on the dissimilarity matrix.

Cluster analysis based on the B_{sim} metric was applied separately to the distribution data at different taxonomic levels (species, genus and family). A separate cluster analysis was implemented using a taxonomically-weighted version of the B_{sim} metric to take into account taxonomic relatedness among sampling units. A taxonomic tree was constructed in BIODIVERSE based on the study data. This tree then served as input for assessing pair-wise dissimilarity between sites while taking into account taxonomic structure (**Figure 5.3**). It is important to note that this taxonomic tree should not be confused with a phylogenetic tree, which provides information on the evolutionary history of species. In this case, the tree is merely a representation of the hierarchical structure between taxonomic levels, but it does highlight relatedness among sites that the analysis at separate taxonomic levels does not.

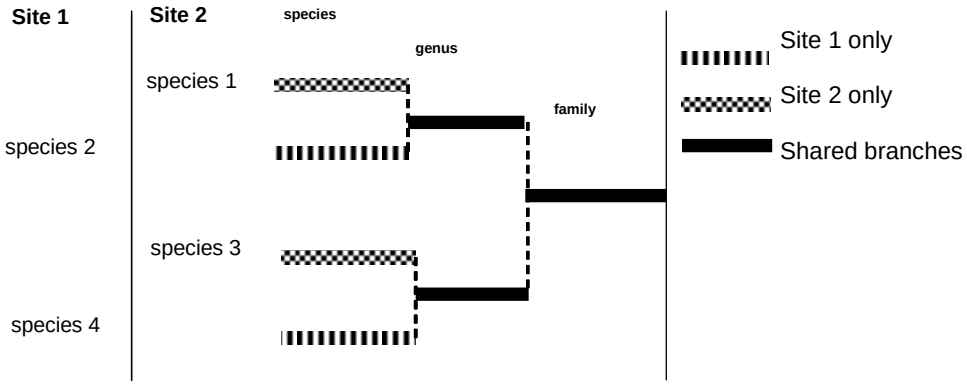


Figure 5.3 Taxonomically-weighted B_{Sim} represents a measure of the contribution of each taxonomic level to the similarity between sites. Instead of focusing on number of species, the dissimilarity value is calculated based on the number of shared branches (solid bar), the number of branches found only in site 1 (vertical bars) and those found only in site 2 (crosshatch).

In order to assess the significance of the resulting dendrograms from the cluster analyses, randomization tests were run in BIODIVERSE for 100 iterations for both the species-level and taxonomically-weighted analyses. The procedure randomized the underlying species data randomly assigning species to different sites but retaining the original species richness of each site. A new cluster analysis was performed at each iteration. Each new randomized dendrogram was assessed for similarity against the original using a Sorensen dissimilarity index (Laffan et al. 2010). BIODIVERSE produces statistics summarizing the mean Sorensen dissimilarity and count of identical nodes for each node (cluster or group) across all iterations. For each node, a low count of identical nodes and high mean Sorensen dissimilarity value indicate the robustness of that node to an alternate random distribution of the species data. This serves as a way of assessing the robustness of the original cluster result (Laffan et al. 2010).

The choice of how many groups to have in a cluster analysis (or at what point to dissect the dendrogram) is still an issue of some debate (Kreft and Jetz 2010). However, my purpose here was not to delineate specific biogeographic regions but to assess the degree of similarity between ecoregions in terms of species composition. For my purposes, I visually determined the number of major splits occurring higher up in the dendrogram. This resulted in a total of sixteen groups. I applied the same number of groups to all resulting dendrograms.

5.2.3.1.2 Cluster Analysis: Grid cells

Cluster analysis was also performed using a gridded approach in order assess species turnover across the region at a finer scale than ecoregions. The same dataset was used but was stratified across the study region into equal-area grid cells at a 1^0 resolution. Each grid cell thus represents a sample site. A resolution of 1^0 or 2^0 (approximately 110-220 km near the equator) is usually considered a reasonable compromise between accuracy and detail for this type of analysis (Kreft and Jetz 2010). For comparison purposes, cluster analysis was performed using both B_{sim} at species level, and taxonomically-weighted B_{sim}.

5.2.3.2 Analyzing Potential Drivers of Species Turnover

Two methods were employed at the ecoregion scale to test the relationship between variation in species turnover and variation in climate (mean annual temperature and annual precipitation). Analysis was carried out at the species level only using the distance matrix (**Figure 5.2**) derived by applying the B_{sim} metric. Non-metric multidimensional scaling and Mantel tests were performed using R v. 2.13.2 (R Development Core Team 2011).

5.2.3.2.1 Non-metric Multidimensional Scaling

A useful method for visualizing resemblance among a set of sites is ordination. Ordination seeks to optimally represent the dissimilarities between sites in a space of low dimensionality. The distances between individuals in ordination space should closely resemble their dissimilarities (Kenkel and Orloci 1986). This is done by finding the best arrangement of individuals in ordination space through an iterative process. Calculated stress values give an indication of how well this arrangement matches original dissimilarities (Kreft and Jetz 2010). Non-metric multidimensional scaling (NMDS) is a form of ordination which basically plots dissimilar objects far apart and similar objects close together. NMDS is recommended because it does not assume normality or linearity of the underlying data (Kreft and Jetz 2010). In addition, NMDS can provide insight about what factors are structuring the data if the ordination axes can be related to environmental gradients. Usually the first axis indicates the main axis of variation and the second indicates residual variation after accounting for the first axis (Kenkel and Orloci 1986).

In this case, NMDS was performed in two dimensions (axes) using the function 'isoMDS' from the MASS library in R v. 2.13.2 (R Development Core Team 2011). Fifty permutations and a tolerance of 0.001 were used to find a stable solution. Spearman's rank correlation tests were performed to test correlations between each NMDS axis and values of mean annual temperature, annual precipitation, longitude and latitude respectively. Climate variables and latitude and longitude were assessed for possible co-linearity in order to help interpret the Spearman test results.

5.2.3.2.2 Mantel Tests

Spatial structuring is often a property of ecological variables and ecosystems (Legendre 1993). This spatial autocorrelation must be taken into account, and statistical tools have been developed to aid ecologists in studying spatially structured variables. For example, Mantel and partial Mantel tests can be used to characterize the relationship between two spatially structured variables (Reynolds and Houle 2002). Mantel tests represent a distance approach used to test the correlation between two distance matrices. Here, Mantel tests were performed to test the correlation between species composition (as measured by the B_{Sim} index) and mean annual temperature, annual precipitation and geographic distance (measured by Euclidean distance). The partial Mantel test can be used for taking into account spatial autocorrelation when testing for the effect of explanatory variables on the response variable (Legendre and Fortin 1989; Smouse et al. 1986). Partial Mantel tests were performed to test the correlation of climatic variables with species turnover while controlling for spatial proximity.

5.3 RESULTS & DISCUSSION

5.3.1 Variation in Species Turnover

The cluster analysis highlighted important patterns of species turnover among ecoregions. Distinct groups of similar ecoregions emerged at all taxonomic levels. My analysis revealed that some ecoregions are more similar to others in species composition. Understanding these differences among ecoregions is important for understanding the spatial structure of tree species diversity across the region.

My analysis notes 16 groups or clusters of ecoregions (**Figures 5.4 - 5.7**). Subsequent splitting along the dendrogram would invariably separate all 34 ecoregions into individual units. However, as stated before, my purpose here is to demonstrate the similarity in species composition among ecoregions, a component which is not taken into consideration in conservation planning when treating ecoregions as equally distinct units. For simplicity, we define the same number of groups (16) when comparing all dendrograms. My results clearly show the hierarchical structure in similarity among ecoregions across all taxonomic levels as evidenced by the dendrograms (**Figures 5.4 - 5.7**).

With some minor variation, the following broad pattern is repeated across all taxonomic levels. Major splits occurring higher up (or to the left of) the dendrogram gave rise to four major groups of ecoregions or 'meta-ecoregions' (**Figures 5.4**). Splits essentially separate four groups consisting of the following: (1) the northern ecoregions located in Mexico including some central ecoregions such as Central American Pine-Oak forest , Central American Dry Forest and Chiapas Depression Dry forest; (2) the ecoregions of the south from Choco-Darien Moist forest to Central American Atlantic Moist forest; (3) the Yucatan ecoregions, including the Peten-Veracruz region; and (4) a group of smaller non-contiguous ecoregions occurring in 'pockets' associated with high altitudes, including Chimalapas Montane forest, Central American Montane forest, Sierra Madre de Chiapas Moist forest, and Chiapas Montane forest.

The resulting taxonomically-weighted dendrogram (**Figures 5.4**) shows that the first split creates two major groups or clusters or ecoregions: (1) a group consisting of most northern ecoregions in Mexico along with Central American Pine-Oak, Montane and Dry forests, and (2) a group consisting of all other ecoregions all the way down to the southern ecoregions in Panama. The next split in group (1) separates out the Chiapas/Chimalapas region along with Central American Montane forests from the rest of the group which consists of mainly Dry, Montane and Pine-Oak forests. In group (2), the next split separates the southern ecoregions along with Central American Atlantic Moist forest from the Yucatan peninsula ecoregions and the moist forests of the Peten-Veracruz region (including Pantanos de Centla and Sierra de los Tuxtlas). The latter set of ecoregions (Yucatan and Peten-Veracruz regions) constitutes the area known as the Maya, Zoque and Olmec forests (or Selva Maya) for which The Nature Conservancy has drafted a regional conservation plan (Pronatura and TNC 2006).

In the taxonomically-weighted analysis, most of the ecoregions in Costa Rica and Panama form a large group, from the Isthmian-Atlantic Moist forest to the Choco-Darien Moist forest (**Figures 5.4**). The Central American Atlantic Moist forest forms another group along with the Misquito Pine forest, while the Central American Pine-Oak forest, Central American Dry forest and Chiapas Depression Dry forest

form another group. Yucatan Moist and Yucatan Dry forests form a separate group while the moist forests of the Peten-Veracruz region form another group. Belizean Pine forest forms a cluster with the Mesoamerican Gulf-Caribbean mangrove ecoregion. Central American Montane forest forms a group with Chimalapas Montane forest. Of the northern Mexican ecoregions, Balsas Dry forest forms a group with Southern Pacific Dry forest. Vercruz Moist and Montane forests cluster with Oaxacan Montane forest. Ecoregions which separate out as individual units at a high level in this dendrogram include: Trans-Mexican Volcanic Belt Pine-Oak forest, Sierra Madre del Sur Pine-Oak forest, Sierra Madre de Oaxaca Pine-Oak forest, Sierra Madre de Chiapas Moist forest and Chiapas Montane Forest.

The species level result (**Figures 5.5**) reveals much of the same structure with the major exception that both Yucatan Moist and Yucatan Dry forests are grouped with the Peten-Veracruz ecoregions. Note that Belizean Pine forest and Mesoamerican Gulf-Caribbean Mangrove are now also grouped with this cluster.

The genus and family level analyses reveal some interesting patterns. For example, at the genus level (**Figures 5.6**), the similarity between Sierra Madre de Chiapas Moist forest and Central American Montane forest and Central American Pine-Oak forest can be seen. At the family level (**Figure 5.7**), the Chiapas Depression Dry forest and Central American Pine-Oak forest seems to bear similarity with Balsas Dry forest.

Another interesting result highlighted by the cluster analysis is that moist forests cluster together with dry forests. For example, the Yucatan Dry and Moist forests form a group, and the Panamanian Dry forest clusters along with other moist forest ecoregions (Isthmian Atlantic and Pacific Moist forest and Choco-Darien Moist forest). This implies that these ecoregions are very similar in species composition but that dry forests contain much of the same species that have adapted to drought conditions (Miles et al. 2006). This has important implications for conservation planning in the face of climate change as it implies that climate change may not lead to significant changes in species composition changes in these ecoregions.

Importantly, some ecoregions which are highlighted across all taxonomic levels as distinct in species composition refer to small ecoregions associated with high elevations, namely: Chimalapas Montane forest and Central American Montane forest (these are grouped together), Sierra Madre de Chiapas Moist forest, and Chiapas Montane forest. My analysis seems to support the uniqueness of the Chiapas/Chimalapas region in representing the confluence of Mexican flora with Central American flora (Olson and Dinerstein 2002). My analysis also seems to support the importance of the Central American Montane forests as pockets of unique flora (Olson and Dinerstein 2002).

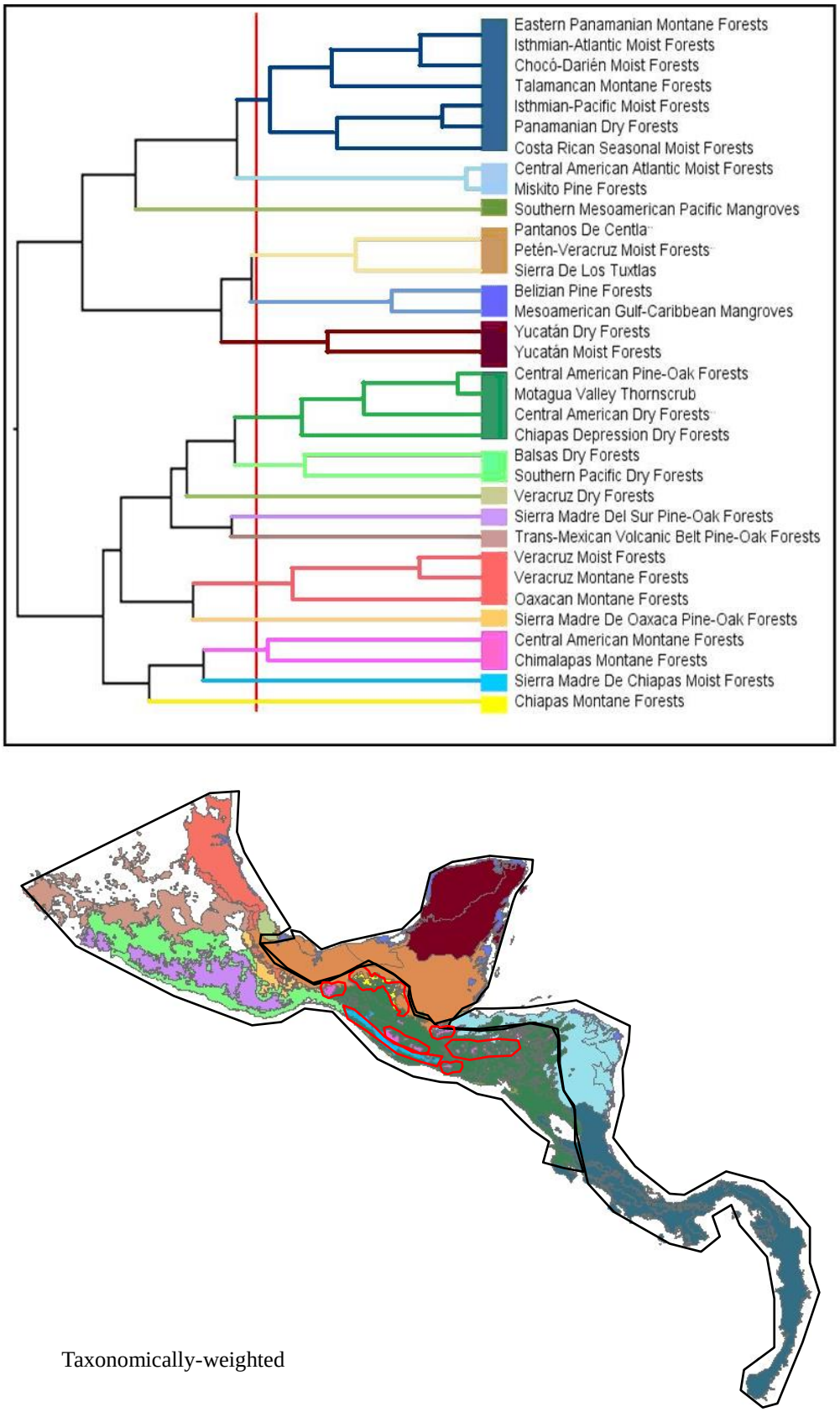


Figure 5.4 Dendrogram and map for cluster analysis of ecoregions based on taxonomically-weighted B_{sim}. Red vertical line on dendrogram shows cut-off point for number of groups. Polygons on map roughly indicate 'meta-ecoregions' based on species composition (red polygons indicate one group of non-contiguous ecoregions that were usually grouped together by the analysis).

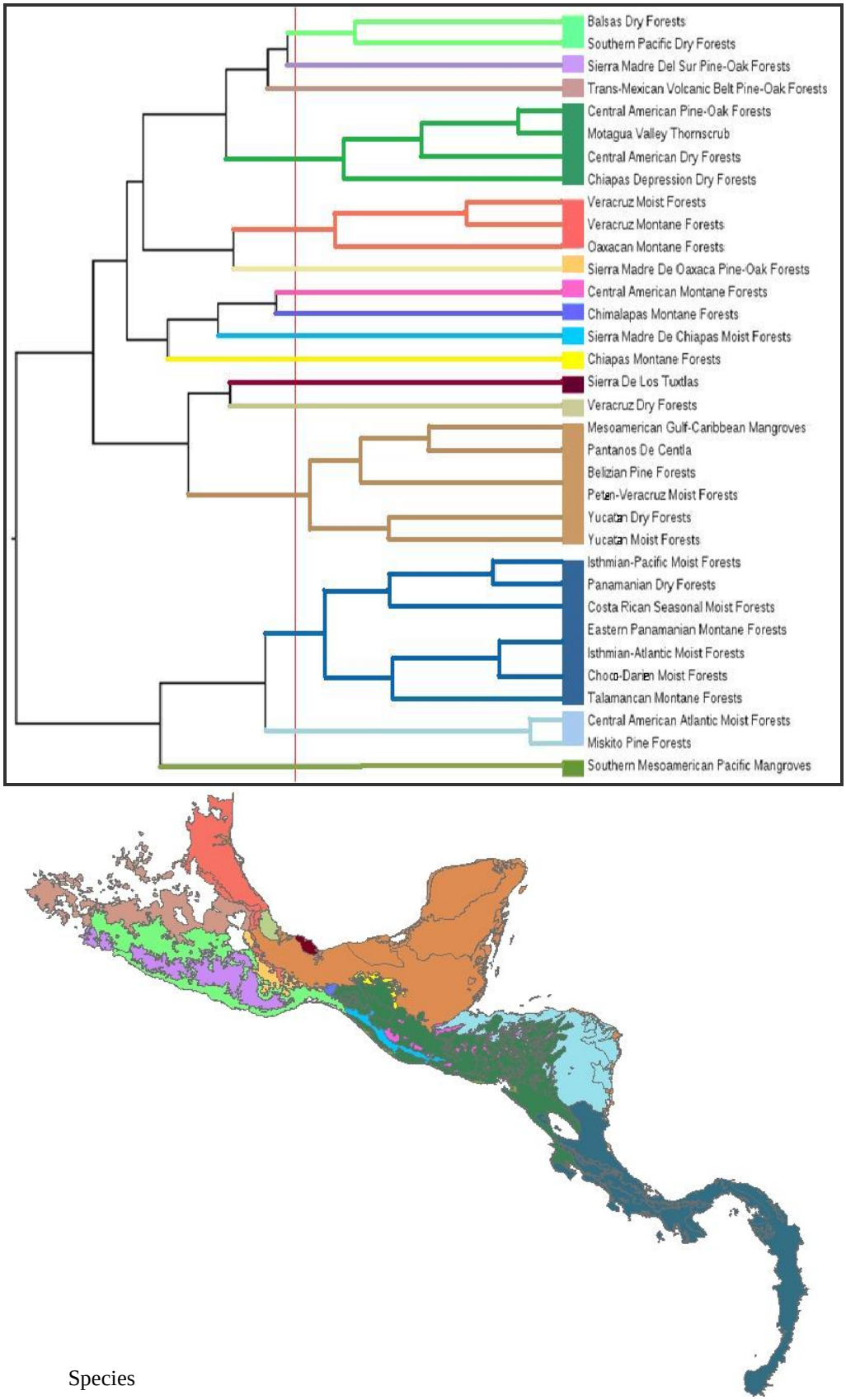


Figure 5.5 Dendrogram and map for cluster analysis of ecoregions based on B_{sim} at species level. Red vertical line on dendrogram shows cut-off point for number of groups.

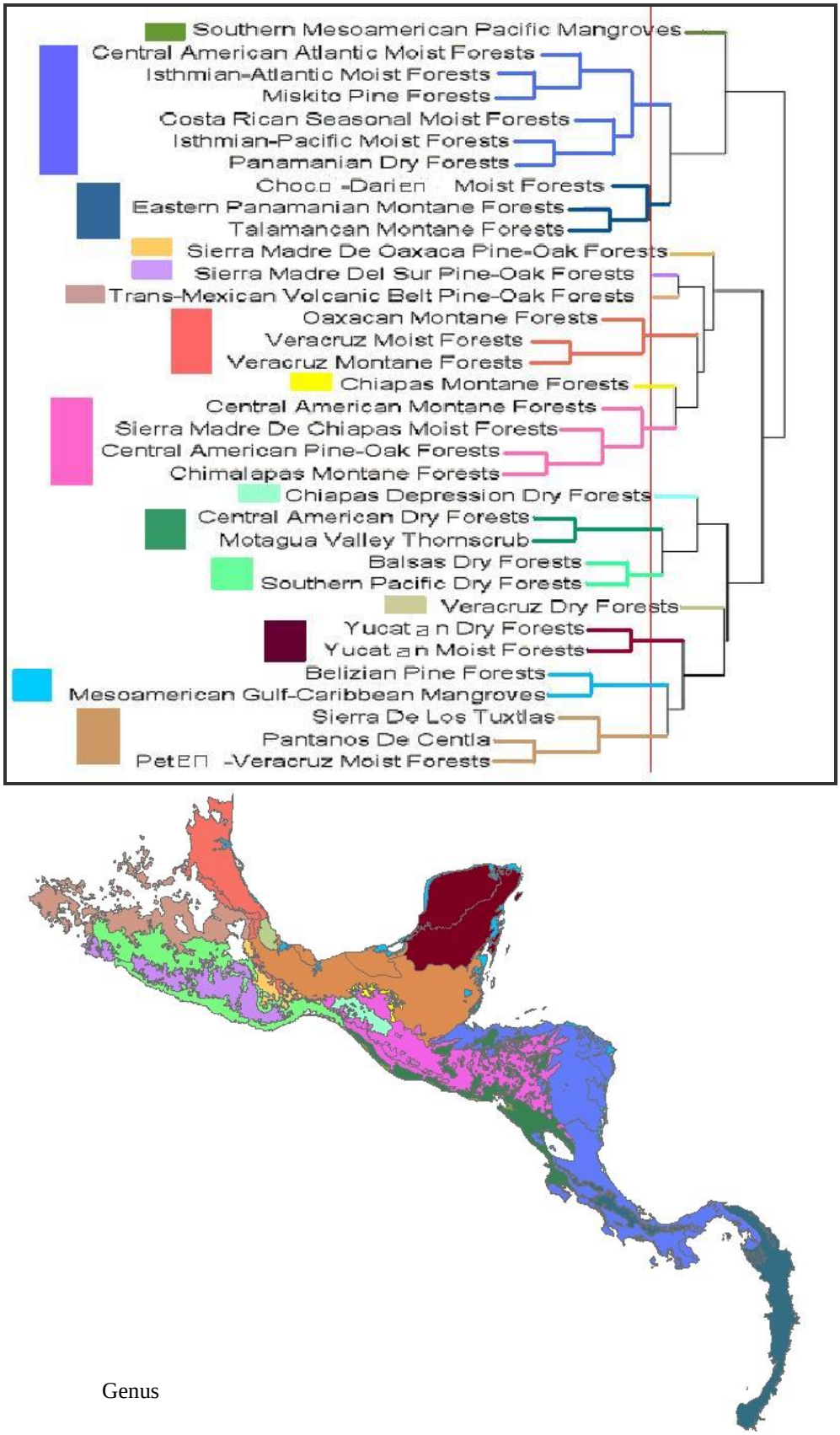


Figure 5.6 Dendrogram and map for cluster analysis of ecoregions based on B_{sim} at genus level. Red vertical line on dendrogram shows cut-off point for number of groups.

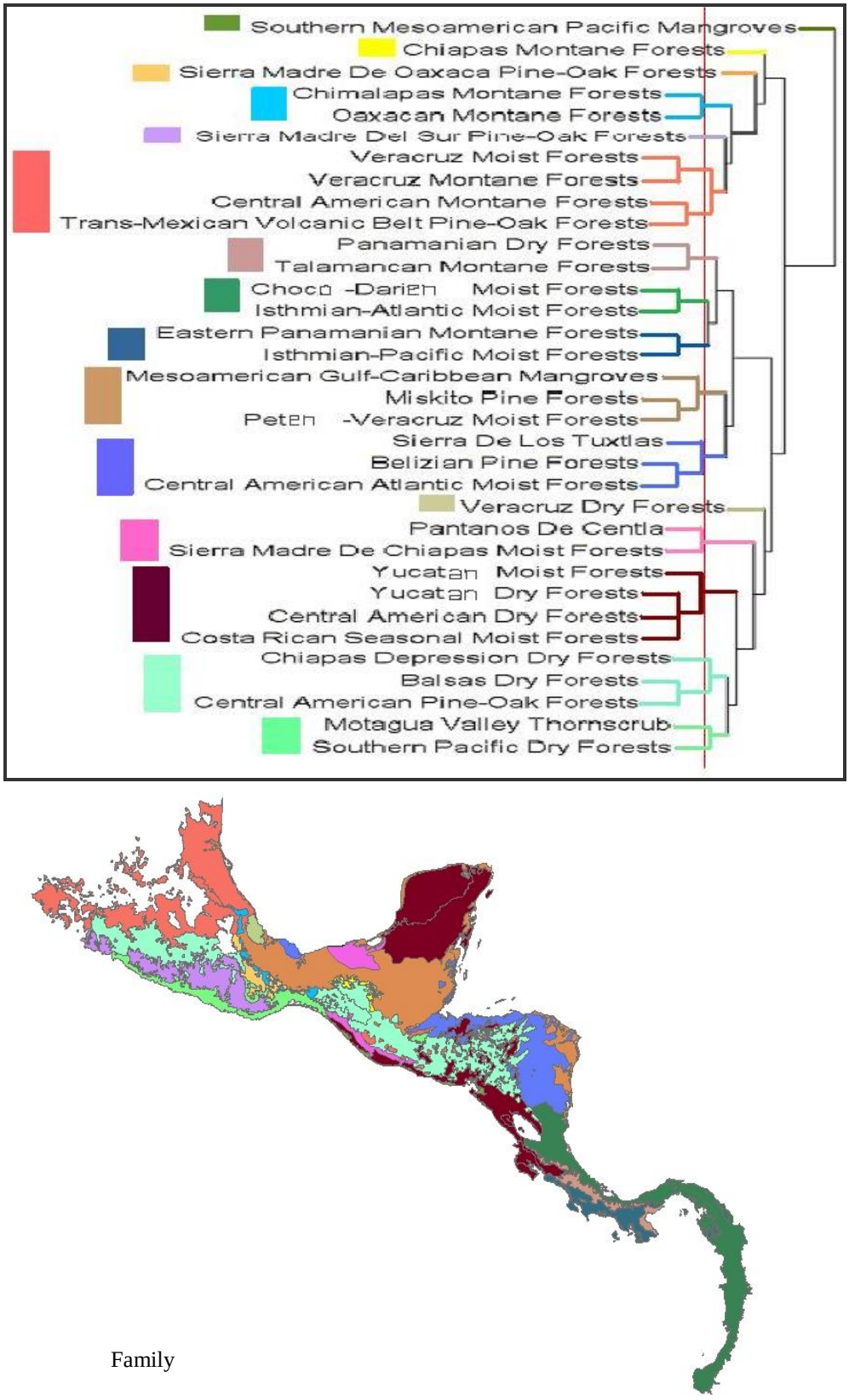


Figure 5.7 Dendrograms and maps for cluster analysis of ecoregions based Bsim at family level. Red vertical line on dendrogram shows cut-off point for number of groups.

The result of the taxonomically-weighted cluster analysis was assessed for robustness using the randomization method described in section 5.2.3.1.1. The mean Sorensen dissimilarity values, count of identical nodes and standard deviations for each node across all randomizations are summarized in **Figure 5.8**. The mean dissimilarity values of all nodes tend to be below 0.5; however, this may be a function of the metric used. The Sorensen metric is sensitive to differences in richness between groups, therefore, the value produced by this metric will be dependent on the number of items (in this case ecoregions) contained within each node (cluster or group). Only two out of 32 nodes (6%) recorded any count identical values, providing us with confidence in the results of the cluster analysis.

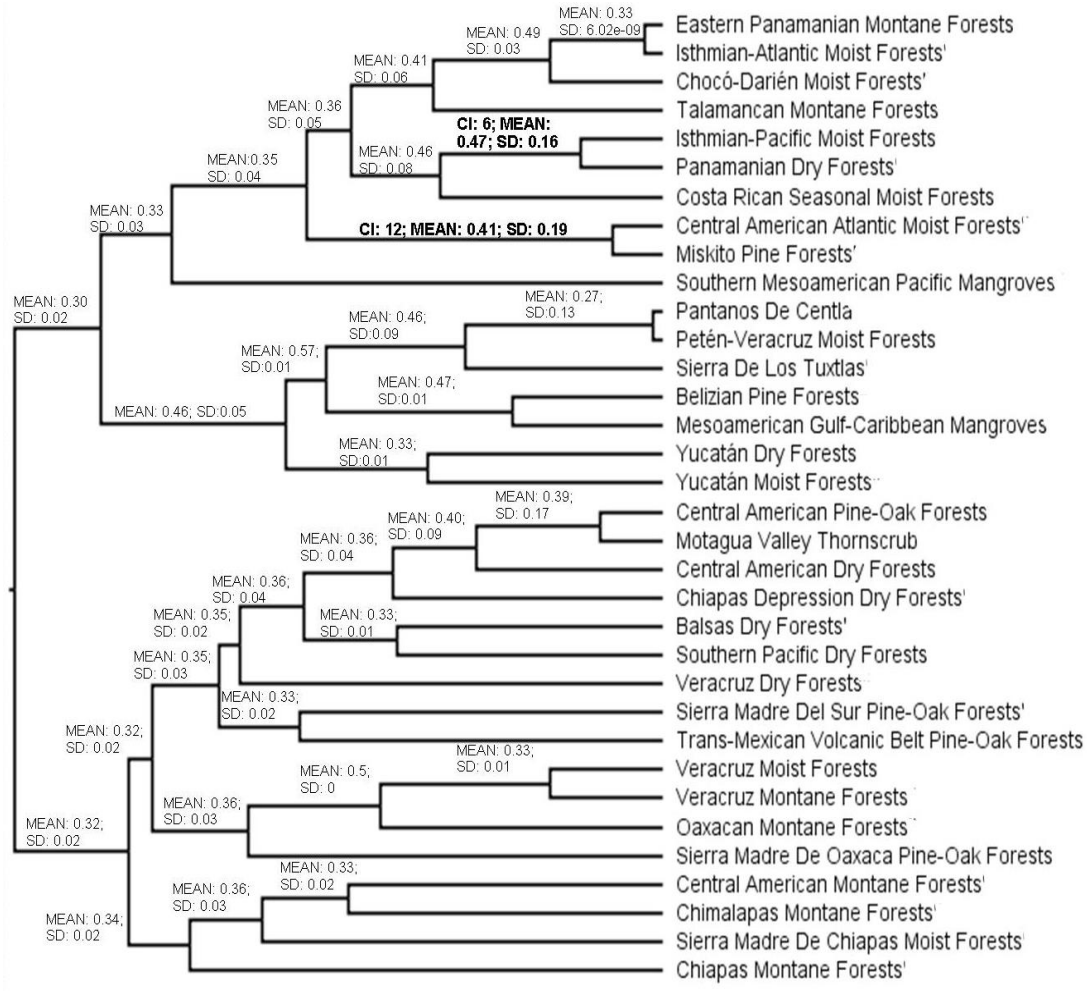


Figure 5.8 Randomization results. Dendrogram summarizing results of randomization test performed on taxonomically-weighted cluster analysis (CI: Count of Identical nodes; S.D.: standard deviation).

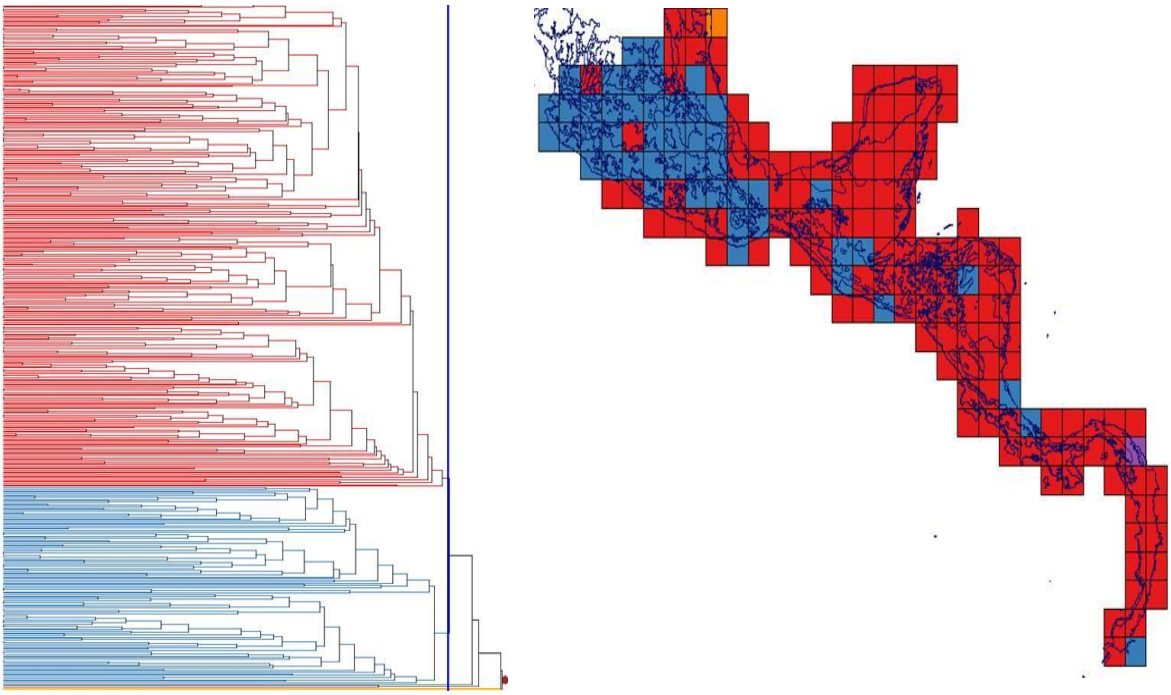
The gridded analysis reveals a similar pattern as that revealed by the ecoregion analysis. Both the taxonomically-weighted and the species level analyses had similar results, so we present the diagrams for only one of these. There is a distinct grouping of grid cells which encompass almost the entire isthmus (red cells in **Figure 5.9a**). Another distinct set of cells occur in the northern area of the study site that corresponds to Mexico (blue cells). Some of these cells reach all the way to the south of the isthmus indicating that they have similar species composition to the cells in Mexico. These include cells encompassing the Central American Pine-Oak and Montane forest ecoregions and the Talamancan Montane forest ecoregion. One cell in the southern region differentiates at the second split in the

dendrogram. This cell encompasses the Eastern Panamanian Montane forest (purple cell in **Figure 5.9a**).

Subsequent splitting up to 13 nodes (**Figure 5.9b**) further separates the above two clusters into three major groups: one group consisting of those cells in the northern ecoregions of Mexico (orange cells), another group consisting of cells bearing close affinity to the south ecoregions of Mexico (red cells) and finally a large group consisting of most grid cells extending from south-central Mexico to the Choco-Darien region (blue cells).

At 13 nodes (**Figure 5.9b**), the analysis shows differentiation of other cells including one encompassing the Central American Pine-Oak ecoregion (pink cell in **Figure 5.9b**), and one in the Choco-Darien Moist forest ecoregion (green cell). These early splits are probably due to scale. At 1^0 resolution, some grid cells are encompassing entire ecoregions such the Eastern Panamanian Montane forest which is less than 3,000 km^2 in area, or portions of ecoregions which are comprised of discontinuous pockets such as the Central American Pine-Oak forest.

a) Six groups



b) Thirteen groups

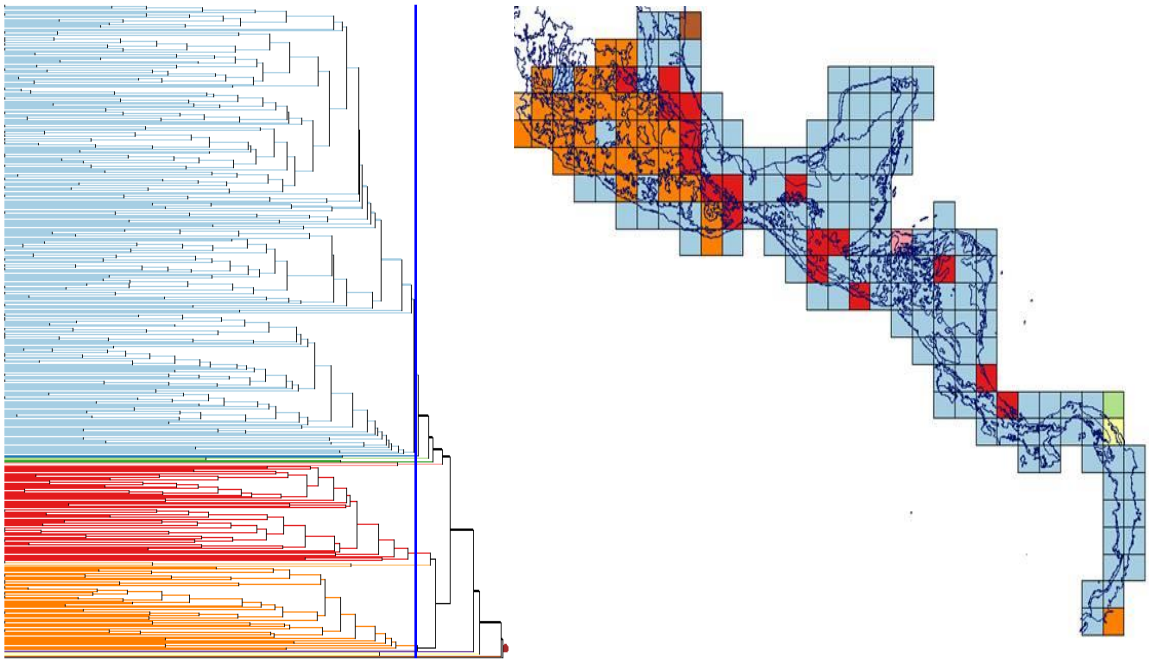


Figure 5.9 Dendrogram and maps for gridded taxonomically-weighted cluster analysis. a) First 6 groups b) first 13 groups.

Cody et al. (2010) use molecular phylogenetic studies to demonstrate that plant lineages had a greater likelihood of dispersing across the Isthmus of Panama region before its closure with South America, that is, before the Great American Biotic Interchange (GABI). This suggests that exchange of flora between South America and Central America has been taking place for longer than previously thought. If this is the case, patterns of species compositional similarity should reflect the main pattern of movement across the Central American isthmus. The results of my gridded analysis of species compositional similarity seem to support previous postulations that the directionality of floral exchange across Central America has been asymmetric from south to north, probably due to this longer history of dispersal opportunity (McCafferty 1998; Vermeij 1991)The overall patterns of similarity seems to support the hypothesis that movement of plants has been mainly northward since or even before (Cody et al. 2010) the GABI.

In addition, the movement of flora north to south through the Central American isthmus is impeded by the presence of the mountain ranges in Mexico. The convergence of the Sierra Madre Occidental and Sierra Madre Oriental, as well as the Sierra Madre del Sur and Trans-Mexican Volcanic Belt (or Sierra Nevada) all present barriers to dispersal. This presents another reason for the asymmetric pattern observed in this and other studies.

In totality, these results seem to suggest that ecoregions which are spatially closer together are more similar in species composition. That is, the variation in species turnover is influenced by geographic distance.

5.3.2 Potential Drivers of Species Turnover

The NMDS performed on the Bsim distance matrix at species level had a stress value of 23% indicating a somewhat stable spatial projection of the dissimilarity matrix into ordination space (Kreft and Jetz 2010). The configuration seemed to confirm the results of the cluster analysis in that ecoregions which were geographically closer together seemed to be clustering together in ordination. However, there also seemed to be an association with annual precipitation (**Figure 5.10**). Many of the Dry Forest ecoregions occurred at the right of the plot while the Moist Forest ecoregions occurred on the left of the plot. Lines representing mean annual temperature and annual precipitation were overlaid over the NMDS projection in order to visualize any trends in these variables. There was a clear pattern of decreasing precipitation going from left to right on the plot. This coincided with the Dry forest ecoregions occurring on the right of the plot.

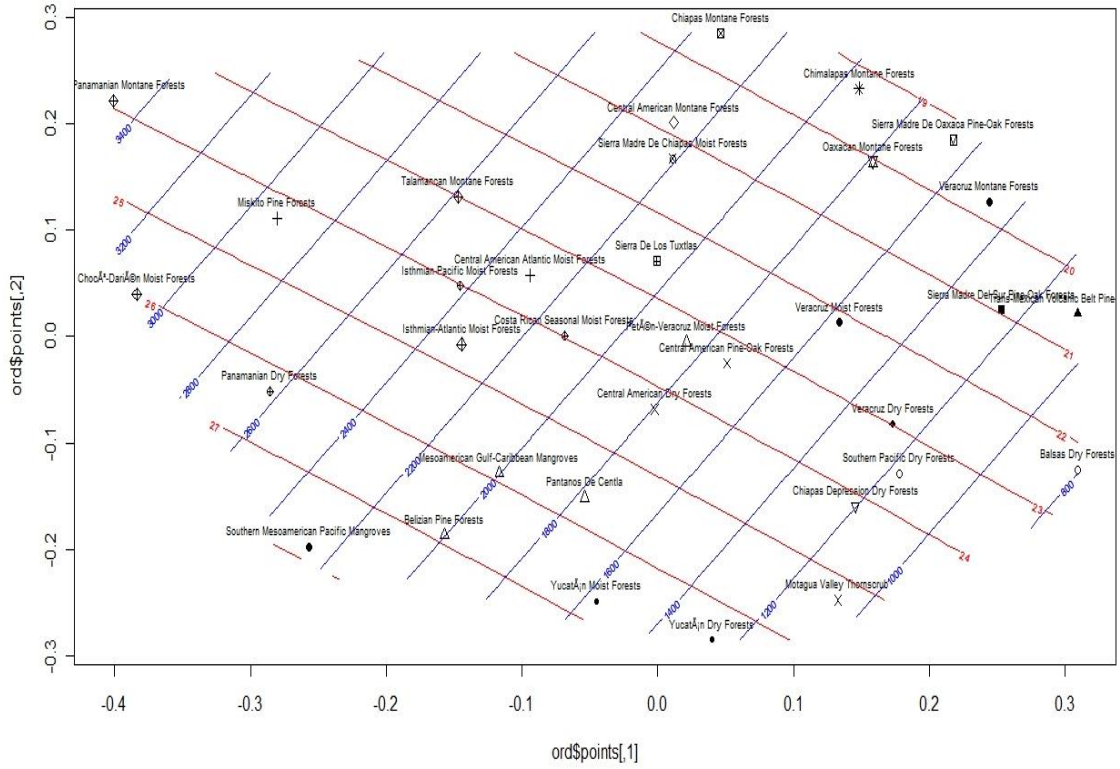


Figure 5.10 Results of NMDS. Red lines represent annual mean temperature and blue lines represent annual precipitation.

I ran correlation tests of these climatic variables, and latitude and longitude against the NMDS axes to determine what factors were structuring the variation in species composition (**Table 5.2**; see **Appendix V** for correlation plots). Results suggest the main axis of variation could be any of the climatic variables as well as geographic distance. Longitude was strongly negatively correlated with the first axis (*rho*=-0.92, *P*<0.001) followed by annual precipitation (*rho*=-0.70, *P*<0.001). Latitude was positively correlated with the first axis (*rho*=0.63, *P*<0.001), while mean annual temperature was negatively correlated (*rho*=-0.51, *P*<0.001). The variable with the strongest negative correlation to the second axis was mean annual temperature (*rho*=-0.62, *P*<0.001). Annual precipitation was positively correlated with the second axis (*rho*=0.49, *P*<0.001).

Despite co-linearity among some of the explanatory variables (**Table 5.3**), the high correlation values obtained in the Spearman tests suggest that geographic distance, annual precipitation and mean temperature are all strongly correlated with variation in species turnover among ecoregions.

Table 5.2 Correlation Test Results. Spearman's Rank Correlation test results for climatic variables and geographic coordinates against NMDS axes.

NMDS axis	Explanatory Variable	Spearman's Rank Correlation *rho*	*P*-value (N=34)
1st axis	Longitude	-0.92	$<2.2x10^{-16}$ *
	Annual Precipitation	-0.70	$6.97x10^{-06}$ *
	Latitude	0.63	0.00011*
	Mean Annual Temperature	-0.51	0.0022*
2nd axis	Mean Annual Temperature	-0.62	0.0001137*
	Annual Precipitation	0.49	0.0003*
	Latitude	-0.228	0.195
	Longitude	-0.031	0.862

* Statistically significant

Table 5.3 Co-linearity results.

Regression variables	R^2	*P*-value
Mean temperature VS		
Latitude	0.007	0.276
Longitude	0.15	0.013*
Annual precipitation VS		
Longitude	0.46	$6.30x10^{-06}$*
Latitude	0.44	$1.04x10^{-05}$*
Meant temperature VS Annual precipitation	-0.01	0.446

* Statistically significant

Results of the Mantel tests support a strong positive correlation between similarity in species composition and geographic distance (**Figure 5.11**). Mantel test results indicate that compositional similarity was most significantly correlated with geographic distance (r=0.48, P<0.01) followed by annual precipitation (r=0.32, P<0.01) and then temperature (r=0.25, P<0.01). Note that the Mantel test provides a measure of the strength of the correlation between distance or similarity matrices. As such it varies between +1 and -1, but the sign itself does not necessarily imply a positive or negative correlation (Reynolds and Houle 2002).

Partial Mantel tests were performed to test the effect of the climatic variables on species turnover while controlling for spatial proximity. In this case, the change in correlation with temperature while controlling for geographic distance was negligible (r=0.28, P<0.01). There was a 0.15 decrease in the correlation with annual precipitation when spatial proximity was taken into account (r=0.17, P<0.01) but the relationship remained significant.

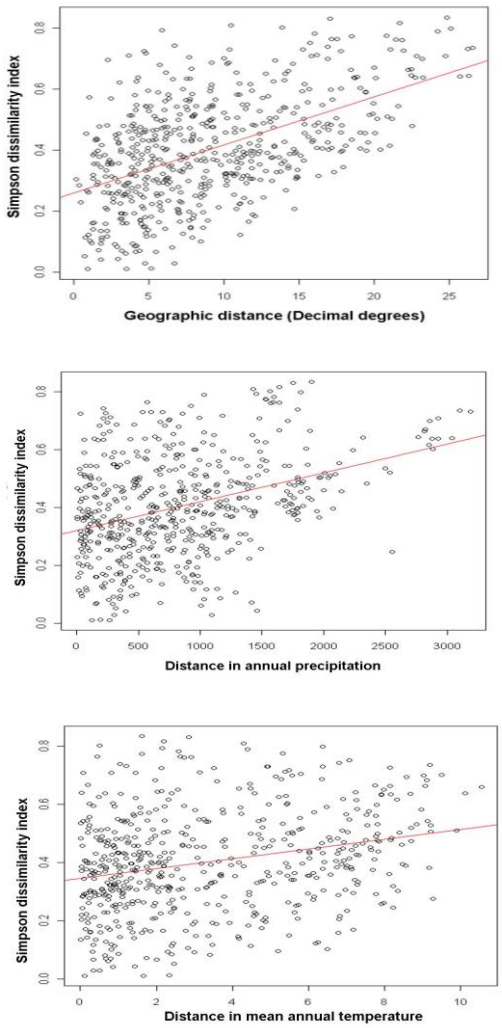


Figure 5.11 Relationship between similarity in species composition of ecoregions and their similarity in a) geographic distance b) annual precipitation c) mean annual temperature. Points are distances among all pairs of ecoregions (N = 34, giving a total of N (N–1) / 2 = 561 pair-wise comparisons).

These results seem to indicate that the variation in species composition between ecoregions may be explained mostly by spatial distance, but that the effects of precipitation and temperature cannot be disregarded. They still seem to be affecting variation in species turnover even when spatial proximity is considered.

5.4 CONCLUSION

I found great similarity at all taxonomic levels among ecoregions commonly defined as equally distinct units of species assemblages. Using species collection records from across the region to assess species turnover, I show that ecoregions cannot be defined as equally different. Many ecoregions share similar species composition. This needs to be taken into account when comparing these ecoregions to others which have more distinct species composition. Conservation effort should be applied cautiously at the ecoregion level bearing in mind patterns of species turnover at this scale.

Based on species turnover, four ‘meta-ecoregions’ could be defined that highlight a) those ecoregions with an affinity to the Mexican regional flora b) those bearing more affinity to the southern flora and flora of South America c) the Yucatan Peninsula ecoregions and d) small ecoregions associated with high elevations. This latter group was highlighted as compositionally distinct across all taxonomic levels and therefore represents regions which should be prioritized as they may harbour species which are not found elsewhere in the Mesomerican region.

Dry forest ecoregions seem to have similar species composition as moist forest ecoregions. I postulate that this has important implications for planning since it suggests that climate change may not necessarily alter the species composition in these areas. Therefore, these areas may represent regions which are more resilient to changing conditions.

Finally, my analysis indicates that patterns of species turnover among ecoregions may be driven primarily by geographic distance. This implies the importance of historical factors (such as the closure of the Panama isthmus) and other factors like dispersal ability or barriers to dispersal. My analysis shows, however, that the effect of climatic variables cannot be discounted since these still seemed to be accounting for some variation in species turnover even when spatial distance was accounted for. It is therefore possible that species turnover patterns among ecoregions reflect the effects of spatial distance and other spatially structured variables like climate.

5.5 REFERENCES

Anderson, M. J., Crist, T. O., Chase, J. M., Vellend, M., Inouye, B. D., Freestone, A. L., Sanders, N. J., Cornell, H. V., Comita, L. S., Davies, K. F., Harrison, S. P., Kraft, N. J. B., Stegen, J. C. and Swenson, N. G., 2011. Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology Letters*, *14* , 19–28.

Barthlott, W., Hostert, A., Kier, G., Kuper, W., Kreft, H., Mutke, J., Daud Rafiqpoor, M. and Henning Sommer, J., 2007. Geographic patterns of vascular plant diversity at continental to global scales. *Erdkunde*, *61* (4), 305–315.

Baselga, A., 2010. Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, *19* (1), 134–143.

Bridgewater, S., Ratter, J. a. and Felipe Ribeiro, J., 2004. Biogeographic patterns, Beta-diversity and dominance in the cerrado biome of Brazil. *Biodiversity and Conservation*, *13* (12), 2295–2317.

Chase, J. M., 2010. Stochastic community assembly causes higher biodiversity in more productive environments. *Science (New York, N.Y.)*, *328* (5984), 1388–1391.

Cody, S., Richardson, J. E., Rull, V., Ellis, C. and Pennington, R. T., 2010. The Great American Biotic Interchange revisited. *Ecography*, 33, 326-332.

Condit, R., Pitman, N., Leigh, E. G., Chave, J., Terborgh, J., Foster, R. B., Núñez, P., Aguilar, S., Valencia, R., Villa, G., Muller-Landau, H. C., Losos, E. and Hubbell, S. P., 2002. Beta-diversity in tropical forest trees. *Science (New York, N.Y.)*, *295* (5555), 666–669.

Giam, X., Sodhi, N. S., Brook, B. W., Tan, H. T. W. and Bradshaw, C. J. a., 2011. Relative need for conservation assessments of vascular plant species among ecoregions. *Journal of Biogeography*, *38* (1), 55–68.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. and Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25* (15), 1965–1978.

Hill, A. W., Otegui, J., Arino, A. H. and Guralnick, R. P., 2010. *GBIF Position Paper on Future Directions and Recommendations for Enhancing Fitness-for-Use Across the GBIF Network*. *Nature*. Copenhagen.

Kenkel, N. C. and Orloci, L., 1986. Applying Metric and Nonmetric Multidimensional Scaling to Ecological Studies: Some New Results. *Ecology*, *67* (4), 919–928.

Kier, G., Mutke, J., Dinerstein, E., Ricketts, T. H., Kuper, W., Kreft, H. and Barthlott, W., 2005. Global patterns of plant diversity and floristic knowledge. *Journal of Biogeography*, *32* (7), 1107–1116.

Kraft, N. J. B., Comita, L. S., Chase, J. M., Sanders, N. J., Swenson, N. G., Crist, T. O., Stegen, J. C., Vellend, M., Boyle, B., Anderson, M. J., Cornell, H. V., Davies, K. F., Freestone, a. L., Inouye, B. D., Harrison, S. P. and Myers, J. a., 2011. Disentangling the Drivers of Diversity Along Latitudinal and Elevational Gradients. *Science*, *333* (6050), 1755–1758.

Kreft, H. and Jetz, W., 2010. A framework for delineating biogeographical regions based on species distributions. *Journal of Biogeography*, *37* , 2029–2053.

Laffan, S. W., Lubarsky, E. and Rosauer, D. F., 2010. Biodiverse, a tool for the spatial analysis of biological and related diversity. *Ecography*, *33* (May), 643–647.

Legendre, P. and Fortin, M.-J., 1989. Spatial Pattern and Ecological Analysis. *Vegetatio*, *80* (2), 107–138.

Legendre, P., 1993. Spatial Autocorrelation : Trouble or New Paradigm ? *Ecology*, *74* (6), 1659–1673.

McCafferty, W. P., 1998. Ephemeroptera and the Great American Interchange. *Journal of the North American Benthological Society*, *17* (1), 1–20.

Mcdonald, R., Mcknight, M., Weiss, D., Selig, E., Oconnor, M., Violin, C. and Moody, a, 2005. Species compositional similarity and ecoregions: Do ecoregion boundaries represent zones of high species turnover? *Biological Conservation*, *126* (1), 24–40.

Miles, L., Newton, A. C., DeFries, R. S., Ravilious, C., May, I., Blyth, S., Kapos, V. and Gordon, J. E., 2006. A global overview of the conservation status of tropical dry forests. *Journal of Biogeography*, *33* (3), 491–505.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'amico, J. a., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P. and Kassem, K. R., 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience*, *51* (11), 933-938.

Olson, D. M. and Dinerstein, E., 1998. The Global 200: A Representation Approach to Conserving the Earth's Most Biologically Valuable Ecoregions. *Conservation Biology*, *12* (3), 502–515.

Olson, D. M. and Dinerstein, E., 2002. The Global 200: Priority Ecoregions for Global Conservation. *Annals of the Missouri Botanical Garden*, *89* (2), 199–224.

Reynolds, C. E. and Houle, G., 2002. Mantel and Partial Mantel Tests Suggest Some Factors That May Control the Local Distribution of Aster laurentianus at Îles De La Madeleine , Québec. *Plant Ecology*, *164* (1), 19–27.

Smouse, P. E., Long, J. C. and Sokal, R. R., 1986. Multiple Regression and Correlation Extensions of the Mantel Test of Matrix Correspondence. *Systematic Zoology*, *35* (4), 627–632.

Vermeij, G. J., 1991. When Biotas Meet: Understanding Biotic Interchange. *Science*, *253* , 1099–1104.

Whittaker, R. J., Araújo, M. B., Jepson, P., Ladle, R. J., Watson, J. E. M. and Willis, K. J., 2005. Conservation Biogeography : assessment and prospect. *Diversity and Distributions*, 3–24.

6 CHAPTER 6

SYNTHESIS & DISCUSSION

The concept of beta diversity is receiving growing research attention, and there is a growing body of literature exploring the application of beta diversity in the context of conservation and climate change planning. Beta diversity is the amount of variation in species composition from one area to another. Knowledge of beta diversity patterns can help inform conservation efforts by helping to plan for persistence of processes that give rise to diversity. Observations of beta diversity patterns can help to identify spatial catalysts—features that can demarcate historical influences on species composition (Pressey et al. 2007). Knowledge of beta diversity patterns can thus make conservation efforts more effective at planning for processes. This thesis aimed to model and to analyze spatial turnover patterns in tree species composition throughout the Mesoamerica hotspot, in order to inform conservation planning efforts. The novelty within this body of research revolves around (i) the taxa and regional scale utilized (ii) the focus on beta diversity patterns instead of the often-cited species richness patterns, and (iii) the integration of beta diversity with other biodiversity assessment approaches to answer questions on species loss and protected areas representativeness within the region.

In this chapter, first I summarize the major research findings linking them to the original objectives and research questions posed by this thesis. I then summarize the major ecological implications arising from this research and place them within the context of ecological theory of diversity patterns. Third, I critically evaluate the methodological approaches utilized in this research, and that form a part of broader issues in ecological modelling. Fourth, I review the major conservation implications and recommendation arising from this thesis. Finally, I list some ideas for future research.

6.1 Summary of findings

6.1.1 Objective 1: Implement the GDM approach to predict patterns of compositional turnover in plant species across the Meoamerican region. (**Chapter 2**)

What variables best predict the variation in tree species composition patterns?

Geographic distance was consistently one of the most important predictors across all the models that were implemented, but its relative importance depended on the beta diversity index used, reflecting whether the index was capturing differences in species richness instead of species replacement. For the B_{sim} index, temperature predictors were more important than predictors based on precipitation, when elevation was excluded from the models. This is in contrast to the long-known correlation between neotropical plant species richness and precipitation (Gentry 1988), suggesting that gradients in temperature are more important in explaining beta diversity patterns than precipitation gradients. This result reflects the variation

in plant adaptations to temperature and in particular adaptations to the narrower climatic niches of tropical mountains (Buckley and Jetz 2008).

How does the choice of a beta diversity index affect the model outcomes?

Deviance explained was higher when the B_{sor} index was used, reflecting the focus of this index in capturing differences in species richness and limiting its utility for detecting species turnover patterns. The B_{sim} index is more conservative in capturing only differences as a result of species replacements (Koleff et al. 2003). The choice of index was also crucial to the relative importance of predictors with the B_{sor} index consistently showing precipitation predictors as the most important, again suggesting it captures differences in species richness (Gentry 1988) (**Appendix II**).

How do tree species compositional patterns vary across the region compared with other regional classifications (e.g. WWF ecoreigons) or turnover patterns of other taxa?

The cluster analysis based on turnover in species composition (using B_{sim} index) gave surprisingly similar results to the WWF-defined ecoregions classification. This is somewhat surprising since a model based on such "gross (*sic*) biophysical features" (Olson et al. 2001) as temperature and precipitation was not expected to show such similar differentiation. There were, however, areas in which the B_{sim} classification showed more differentiation than ecoregions, and these may warrant more investigation, such as the Peten-Veracruz Moist Forests ecoregion. Congruent patterns of high beta diversity and richness in trees and other taxa were located at the north, south and within the centre of the study region, encompassing the complex of dry, montane and pine-oak forests of the region.

6.1.2 Objective 2: Apply the GLOBIO framework and GDM approach to assess the impact of land use change on tree species diversity. (**Chapter 3**)

What proportion of habitat and species has been lost as a result of land use changes up to the year 2000?

Mean loss of effective habitat area (EHA) across the region was estimated at 46.7%, resulting in estimated mean species loss of 14.3%. This is a staggering figure for habitat loss but not outside the range of previous studies that only looked at loss of vegetation cover and did not incorporate beta diversity. The estimated figure for tree species loss was also comparable to IUCN Red List data from 1998 that indicated 19% of regional species were threatened or extinct.

What is the effect of land use change on both areas that are compositionally more similar to other areas in the region (low beta diversity), and those that are less similar (high beta diversity)?

Land use change has led to high levels of EHA loss in both areas of high beta diversity and low beta diversity. There was only 3.3% difference in estimated species loss between ecoregions. Ecoregions with low beta diversity were estimated to have lost only marginally less of their original species.

How does habitat and species loss vary among ecoregions?

Southern ecoregions had lowest REHA suggesting that these regions had lost more of their original species. However, the other ecoregions had also lost substantial original species as the variation in REHA and estimated species loss showed little variation among ecoregions. The ecoregion with the lowest REHA in this analysis was Isthmian Pacific Moist Forest with 50.9% REHA. Remaining habitat cover of 10-30% (i.e. habitat loss of 70-90%) suggests negative effects on species persistence in the form of significant decreases in woody species richness (de Albuquerque and Rueda 2010), and higher probability of population extinction for species (Lima and Mariano-Neto 2014). Although the lowest REHA in this study was higher than this range of 10-30%, a relatively modest further reduction in habitat of 20% would push this ecoregion nearer to the threshold, below which species persistence is negatively affected by low vegetation cover.

6.1.3 Objective 3: Assess the representativeness of the regional protected areas system in relation to species composition patterns. (**Chapter 4**)

6.1.3.1 How representative of tree diversity is the PA system as a whole?

Protected areas are the primary device in a conservationist's tool box and in the overall effort to stem habitat and species loss. Evidence suggests that protected areas reduce the rate of loss of natural habitat within their boundaries, but there is variation in their effectiveness and protected areas in regions such as the tropics continue to experience habitat loss (Gaston et al. 2008). Unlike other regional assessments of proportion of protected habitat that focused only on total area within protected areas (Soutullo et al. 2007; Jenkins and Joppa 2009), this analysis is novel in that it incorporated similarity (beta diversity) between and among protected and unprotected cells. The estimates of proportion protected therefore represent estimates of the proportion of similar habitat protected for all grids cells across the entire region. Although this approach has been implemented globally (Ferrier et al. 2004), this regional analysis provides much needed insight about the variation in protection at a finer scale.

Across the entire region, the 2010 protected area system was estimated to protect a mean of 22.4% of original habitat based on the GDM model at the assumed pristine condition. Estimates of protection almost doubled after habitat loss up to the year 2000 was incorporated into the analysis. This doubling of protection resulted from the reduction in total effective habitat area (EHA) available for original species due to habitat loss up to year 2000. These results reiterate the observation that protected areas protect habitat within them, but that species loss continues outside of protected areas, resulting in increasing isolation of protected areas (Naughton-Treves et al. 2005). The fact that the protection estimates presented in this chapter refer to similar

habitat for original species, suggests that the habitat for these species is increasingly shrinking and being found only within isolated protected areas. This could result in further species loss if distances to such areas are prohibitive to dispersal.

The areas richest in endemic plant species within hotspots such as Mesoamerica are said to have the least habitat remaining and protected (Pimm and Raven 2000). At this regional level, some areas with high dissimilarity (and therefore more likely to harbour narrow-range or endemic species) had low REHA (remaining EHA) in concordance with Pimm and Raven (2000), but had high protection. Additionally, some areas of high dissimilarity were shown to have high REHA but low protection e.g. ecoregions in Mexico. These results highlight the usefulness of regional analyses in unmasking patterns of remaining habitat and protection that are different to globally generalized patterns.

6.1.3.2 What ecoregions are best or least represented in the PA system?

There was a clear gradient across the region in proportion protected. Ecoregions at the south were the most protected, with ecoregions showing progressively less protection to the north. About half of ecoregions with high dissimilarity had below average protection, and half had above average protection. Moist forest ecoregions were more represented in the protected area system than dry forest ecoregions, and all pine-oak forest ecoregions had below average protection. Ecoregions with high dissimilarity and low protection were mostly those ecoregions in Mexico. The clear asymmetries identified in this analysis suggest that conservation efforts have been mostly targeted at one ecoregion type, moist forest ecoregions, and thus more of the same is being protected.

It is important to note that this analysis did not incorporate degradation within protected areas that may not necessarily lead to changes in forest cover, such as over-exploitation of NTFPs (non-timber forest products) in the case of tree taxa. The empty forest syndrome (Redford 1992) need not apply only to vertebrate taxa, since forests provide a suite of NTFPs, such as edible and medicinal plants, and resources for artisanal work (e.g. basket weaving). Furthermore, depletion of other species in the forest system, such as from hunting, can lead to cascading declines in other taxa (Wilkie et al. 2011).

6.1.3.3 What environmental factors are associated with level of protection within the protected area system?

Results of this analysis revealed the bias towards protecting moist forest ecoregions. The results confirm previous analyses showing negative correlation between elevation and level of protection in Mexico and a positive correlation in these two variables in Costa Rica and Panama (Joppa and Pfaff 2009). The results presented here were in concordance with those finding. In addition, results suggest that ecoregions with high dissimilarity (low EHA) are the least protected.

6.1.4 Objective 4: Evaluate whether the ecoregions classification system captures similarity in species composition (species turnover or beta diversity) among ecoregions.

How does species compositional similarity vary among ecoregions?

Four 'meta' ecoregions were distinguished mostly separating the southern ecoregions from the northern ecoregions, and differentiating the Yucatan peninsula and moist forests of the Altantic coast in the region. One of the resulting 'meta' ecoregions comprised the montane forests of Central American and the Chiapas region. Results seemed to provide evidence for the asymmetric dispersal of flora from south to north (McCafferty 1998; Vermeij 1991), possibly as a result of longer dispersal opportunities even before the GABI (Cody et al. 2010).

What is the contribution of spatial distance, mean annual temperature and mean annual precipitation as potential drivers of variation in species turnover among ecoregions?

Results of NMDS analyses and Mantel and partial-Mantel tests suggest that variation in compositional similarity among ecoregions could be mostly explained by geographic distance, followed by differences in annual precipitation, and then temperature. Although spatial distance was the strongest correlate to beta diversity, the effects of annual precipitation and mean temperature could not be disregarded.

6.2 Ecological implications of the research

Varying thresholds of habitat loss have been linked to negative effects on species diversity through such mechanisms as decreases in species persistence, higher rates of genetic drift (Lima and Mariano-Neto 2014), and decreases in woody species richness (de Albuquerque and Rueda 2010). In addition, species loss has been linked to loss of ecosystem services such as carbon storage capacity (Bunker et al. 2005). Efforts to elucidate the relationship between biodiversity and ecosystem function have been severely limited by a lack of studies in highly diverse systems like tropical forests, and by a common focus on species richness rather than species composition (Lewis 2009). Measures of species richness do not adequately capture changes to community structure such as taxonomic homogenization.

Decreased EHA implies loss of species and could signal taxonomic homogenization. Taxonomic homogenization (TH) refers to a temporal decline in beta diversity as a result of an increase in the taxonomic similarity of two or more biotas over time (Olden 2006). The decrease in EHA after land use change indicated that habitat for original species had shrunk from the assumed pristine condition until the year 2000. The areas of low EHA in the pristine condition (**Chapter 2**) had expanded in the year 2000 leading to reduced EHA across the entire region (**Chapter 3**). The effect of habitat loss for originally occurring species is particularly important in the case of areas having narrow-ranged species (such as endemics) where beta diversity is likely to be high (McKnight et al. 2007). In this analysis, decreased EHA at the year 2000 could represent TH since the similarity between grid cells is taken into account when calculating EHA. However,

quantifying TH requires information on the identity of the species at the two different time periods (Olden 2006), which the GDM-MSA analysis (**Chapter 3**) does not implement since the similarity between cells remained constant and only the condition (MSA) changed between the time periods. However, other studies have shown that species richness loss lags taxonomic homogenization. Woodland plant communities did not exhibit significant changes in species richness despite loss of beta diversity and loss of species from the regional species pool (Keith et al. 2009). Quantifying effects of habitat loss on diversity is usually done separately through the use of the SAR to estimate loss of species richness and through separate analysis of community composition (Banks-Leite et al. 2012). **Chapter 3** integrates both these approaches and shows how beta diversity can be incorporated to reveal loss of diversity masked by lack of changes in species richness.

The pattern of regional classification based on B_{sim} (**Chapter 2**) was broadly similar to the WWF ecoregional classification. This suggests that the final model presented here is a model of wide-spread species, since it is revealing such large scale patterns. Wide-ranging species have been found to drive patterns of species richness (Buckley and Jetz 2008). The present results suggest that, at this regional scale, beta diversity patterns also seem to be driven by wide-spread species. There is evidence to suggest that tropical forests are not much different from temperate forests in that they are dominated by a few common species concentrated in a few higher taxa (Pitman et al. 2001). Pitman et al. (2001) found "oligarchies" of species in Ecuador and Peru where most trees were found to belong to a set of common species. Similarly, in the Amazon, 227 species of hyperdominant species were found to make up half of all individual trees (ter Steege et al. 2013). Wide-spread species are thought to influence patterns of beta diversity and species richness because they show stronger responses to simple environmental variables (Lennon et al. 2011) or patterns of mean environmental variation (Gaston et al. 2007). The model presented here is based largely on mean climatic variables and thus seems to provide evidence for that hypothesis. In addition, at the regional scale, as in this analysis, species diversity is largely shaped by abiotic, extrinsic, broad-scale factors such as climate, whereas biotic factors such as dispersal ability and competition seem to operate at local scales and over short time periods (Antonelli and Sanmartín 2011; Benton 2009; Field et al. 2009). The results presented (**Chapter 2**) seem to capture only large-scale patterns in beta diversity driven by the responses of wide-spread species to the broad climatic predictors that were utilized.

It would be erroneous to imply, however, that wide-spread or common species are unimportant. In fact, there is growing recognition that common species are the main victims of habitat loss, the main providers of resources and therefore the main subjects of over-exploitation (Gaston 2010). Loss of common species cascades into losses of other species since common species shape their environment e.g. by being involved in large numbers of biotic interactions with other species (Gaston 2010). Thus, even broad-scale patterns have utility in conservation planning due to the interconnectedness of the natural world.

6.3 Evaluation of methodological approaches

6.3.1 Herbarium record data

The problems with using herbarium records for assessing patterns of species diversity have previously been documented. For example, it is assumed that records have been collected randomly and are unbiased with respect to location and environmental gradients (Loiselle et al. 2007). Both these assumptions were found to be violated in studies evalutating the use of herbarium records, with collections found to be poorly representative of climatic gradients and having 'roadside' bias, i.e. the majority of collections occurred near roads and easy access points (Loiselle et al. 2007; Tobler et al. 2007). In addition, species diversity was found to be sensitve to collection effort (Tobler et al. 2007). In the tropics specifically, lack of data is still an issue, suggesting that the systematic collection of survey data is still needed (Cayuela et al. 2009) in addition to data quality control, such as the standardization of plant names (Cayuela et al. 2012). In this thesis, I tried to address some of these issues by utilizing standardized names for the species data (Cayuela et al. 2012), and by sub-sampling the data.

6.3.2 Predictors used

One of the limitations of the model presented in this thesis is that it only considers one mechanism of neotropical diversity and that is abiotic climatic factors. Other contributors to neotropic plant species diversity include time, and biotic factors such as dispersal ability, competition, soil adaptations, and niche conservatism (Antonelli and Sanmartín 2011). Further research could look at implementing more mechanistic models that explicitly incorporate the processes believed to be affecting observed empirical patterns (Gotelli et al. 2009) by (i) including biotic predictors such as soil or dispersal ability, and (ii) explicitly incorporating historical variables such as time since glaciations (Svenning et al. 2010). However, it may be more difficult to implement such mechanistic models at large scales such as this study region, than at smaller scales. For instance, the availability of information on tree species' dispersal ability would limit studies at both spatial scale and taxonomic level. Perhaps only a suite of species would be available for analysis depending on the availability of dispersal ability information.

6.3.3 Sampling method

Although no decline was noted in model performance when training a GDM model with a lower number of site pairs than that used here (Rosauer et al. 2014), that study used phylogenetic beta diversity and was conducted on one amphibian species. It is possible that such a threshold is not transferable to other taxa such as trees and that, as a consequence, the present tree species data set was underutilized. I attempted to address this problem by sub-sampling the data and running GDM models on these different data samples (**Chapter 2**). However, I only sampled the data three times and ran the model with these three data sets to compare the results. It is possible that more sub-samples would have been necessary; however, computational limitations made this unviable. An alternative sampling method would be to sample site-pairs instead of sites. The

disadvantage of such an approach, however, is that each site is not compared to every other site in the sample.

Another possible shortcoming in the sampling of data could be the fact that grid cells with five or more species were used to train the model. This could be interpreted as a low threshold. However, it is important to note that the sparse nature of available data throughout the neotropics is one of the motivations for modelling using the GDM technique. GDM is an extrapolative technique that seeks to model turnover in areas where it cannot be accurately calculated using empirical data. The cells containing five or more species were used to calibrate a model that used environmental variables and geographical coordinates to predict differences between sites. Although the inclusion of cells with low numbers of species may increase the stochastic element in the data used when fitting the model, it was hoped that the technique would prove sufficiently robust to enable the underlying deterministic component of species turnover to be calibrated.

Not withstanding these issues with the data, the results from these analyses showed similar patterns that have been previously documented in other studies. For example, the results on estimated habitat loss and species loss were similar to estimates documented in other studies (**Chapter 2 & 3**). Similarly, the analysis assessing protected areas representativeness confirmed results from other studies that have documented biases in the location of protected areas which benefit mainly moist forest habitat, and shown the gap in protecting areas of high species turnover, narrow-ranged species and endemism (**Chapter 4)**. Overall, these similarities seem to confirm that the methods utilized throughout the thesis have produced results robust to interpretation.

6.4 Conservation implications & recommendations

Despite the limitations of the methodological approaches discussed above, important conservation implications can be gleaned from the results of this research. Southern ecoregions had lost most of their original habitat and species up to 2000 (**Chapter 3**) but had the most habitat represented within the current protected areas system (**Chapter 4**). In contrast, northern ecoregions lost marginally less habitat and species up to 2000 (only 8.5% and 3.4% less, respectively) but were substantially less protected by the current protected area system (15.6% less protection). The Central American complex of montane, pine-oak and dry forests and most Chiapas ecoregions were also highlighted as having lost more than 12% of original species, and having low representation in the protected area system. According to the CBD's Global Strategy for Plant Conservation target of 15% representation for each ecological region (Pimm et al. 2014), most of the northern ecoregions are already below this threshold or only marginally above it. One issue raised by such observations is whether the establishment of protected areas mostly in moist forest ecoregions has resulted in "leakage" or displacement of disruption activities to other ecoregion types such as dry forests. Nevertheless, protected areas are still a useful tool in conservation as refuges for rare species populations, and stepping stones of suitable habitat for others (Thomas and Gillingham 2015). While the results of this analysis cannot ascertain displacement, the magnitude of EHA loss across the region together with the strong spatial asymmetry in level of protection and type of ecoregion most protected (moist forests) highlight the need for

biodiversity conservation efforts to be focused on (i) thus far neglected ecoregions, and (ii) mitigating the effects of disruptive activities outside protected areas that also have an effect within them (Laurance et al. 2012). The ecoregions highlighted here, showing congruent spatial patterns of low EHA (high beta diversity) for trees and high beta diversity in other taxa (**Chapter 2**), provide a good starting point for targeting conservation efforts in the Mesoamerican region.

The application of extrapolative methods such as GDM to assess patterns of beta diversity can contribute to the prioritization of conservation efforts. GDM is an extrapolative technique that seeks to predict differences in species composition between sites based on environmental variables and geographical distance. Planning schemes that focus only on species richness measures may be ignoring important aspects of how species composition varies across landscapes. The main output from the GDM modelling process is the EHA (effective habitat area) measure, which is a unitless value that represents the proportion of habitat within a landscape that is similar to a particular site. Such a measure provides a complementary method to species richness for assessing sites for conservation. For example, in estimating species loss from habitat loss, the EHA measure takes into account that a species may be lost at one site but may persist in another site of similar habitat. Estimates of species loss based only on number of species per area (species area relationship) do not account for this effect and may overestimate species loss. Additionally, since EHA is based on similar habitat, it provides an easy was to predict truly distinct sites, that is, those sites that have low habitat similarity with other sites across the landscape. Finally, these predicted distinct sites may indicate areas of endemism. If such predictions differ from accepted knowledge, model predictions could be used to identify new survey sites thereby increasing the knowledge base on potentially narrow niche habitats.

Specific recommendations arising from this research include:

- The need to incorporate beta diversity patterns into conservation planning as species richness lags beta diversity loss.
- The need to target conservation efforts in areas of congruence in high beta diversity for multiple taxa
- The need for more data collection and data standardization in the Neotropics
- The need to implement more mechanistic models that can inform about processes not just patterns

6.5 CONCLUSION & FURTHER RESEARCH

The aim of this thesis to model and analyze spatial turnover patterns in tree diversity was attained. It was revealed that broad-scale patterns of tree species turnover could be obtained from a purely climatic model. The resulting patterns were broadly analogous to the WWF ecoregional classification although some differences could be identified. Habitat loss and species loss across the Mesoamerican region were

quantified and the representativeness of the current protected areas system was assessed with respect to beta diversity patterns.

This work highlighted some avenues of further research, including:

- The need to gather and analyze more quantitative plot-based inventory data. Inventory data with good coverage of the region would provide a complementary route to study beta diversity patterns possibly without the need of extrapolative methods such as GDM
- Assessing congruence in spatial patterns of high beta diversity and high carbon storage potential as measured by biomass
- Exploring beta diversity at the phylogenetic level rather than just the species level
- Conducing beta diversity assessments at smaller scales.

6.6 REFERENCES

Antonelli, A. and Sanmartín, I., 2011. Why are there so many plant species in the Neotropics ? *Taxon*, *60* (2), 403–414.

Banks-Leite, C., Ewers, R. and Metzger, J., 2012. Unraveling the drivers of community dissimilarity and species extinction in fragmented landscapes. *Ecology*, *93* (12), 2560–2569.

Benton, M. J., 2009. The Red Queen and the Court Jester : Species Diversity and the Role of Biotic and Abiotic Factors Through Time. *Science*, *323* (February), 728–732.

Buckley, L. B. and Jetz, W., 2008. Linking global turnover of species and environments. *Proceedings of the National Academy of Sciences of the United States of America*, *105* (46), 17836–17841.

Bunker, D. E., Declerck, F., Bradford, J. C., Colwell, R. K., Perfecto, I., Phillips, O. L., Sankaran, M. and Naeem, S., 2005. Species loss and aboveground carbon storage in a tropical forest. *Science (New York, N.Y.)*, *310* (5750), 1029–1031.

Cayuela, L., Golicher, D. J., Newton, A. C., Kolb, M., de Alburquerque, F. S., Arets, E. J. M. M., M., A. J. R. and Pérez, A. M., 2009. Species distribution modeling in the tropics: problems, potentialities, and the role of biological data for effective species conservation. *Tropical Conservation Science*, *2* (3), 319–352.

Cayuela, L., Granzow-de la Cerda, Í., Albuquerque, F. S. and Golicher, D. J., 2012. Taxonstand: An r package for species names standardisation in vegetation databases. *Methods in Ecology and Evolution*, *3* (6), 1078–1083.

Cody, S., Richardson, J. E., Rull, V., Ellis, C. and Pennington, R. T., 2010. The Great American Biotic Interchange revisited. *Ecography*, *33*, 326-332.

De Albuquerque, F. S. and Rueda, M., 2010. Forest loss and fragmentation effects on woody plant species richness in Great Britain. *Forest Ecology and Management*, *260* (4), 472–479.

Ferrier, S., Powell, G. V. N., Richardson, K. S., Manion, G., M, J., Allnutt, T. F., Cameron, S. E., Mantle, K., Burgess, N. D., Faith, D. P., Lamoreux, J. F., Kier, G., Hijmans, R. J., Funk, V. A., Cassis, G. A., Fisher, L., Flemons, P., Lees, D., Lovett, J. C., Rompaey, R. S. A. R. Van, Overton, J. M., Allnutt, F., Fisher, B. L., Lovett, J. O. N. C. and Rompaey, R. S. A. R. V. A. N., 2004. Mapping More of Terrestrial Global Biodiversity for Conservation Assessment. *BioScience*, *54* (12), 1101–1109.

Field, R., Hawkins, B. a., Cornell, H. V., Currie, D. J., Diniz-Filho, J. A. F., Guégan, J.-F., Kaufman, D. M., Kerr, J. T., Mittelbach, G. G., Oberdorff, T., O'Brien, E. M. and Turner, J. R. G., 2009. Spatial species-richness gradients across scales: a meta-analysis. *Journal of Biogeography*, *36* (1), 132–147.

Gaston, K. J., Davies, R. G., Orme, C. D. L., Olson, V. A., Thomas, G. H., Ding, T.-S., Rasmussen, P. C., Lennon, J. J., Bennett, P. M., Owens, I. P. F. and Blackburn, T. M., 2007. Spatial turnover in the global avifauna. *Proceedings. Biological Sciences / The Royal Society*, *274* (1618), 1567–74.

Gaston, K. J., Jackson, S. F., Cantú-Salazar, L. and Cruz-Piñón, G., 2008. The Ecological Performance of Protected Areas. *Annual Review of Ecology, Evolution, and Systematics*, *39* (1), 93–113.

Gaston, K. J., 2010. Valuing common species. *Science*, *327* (5962), 154–155.

Gentry, A. H., 1988. Changes in Plant Community Diversity and Floristic Composition on Environmental and Geographic Gradients. *Annals of the Missouri Botanical Garden*, *75* (1), 1–34.

Gotelli, N. J., Anderson, M. J., Arita, H. T., Chao, A., Colwell, R. K., Connolly, S. R., Currie, D. J., Dunn, R. R., Graves, G. R., Green, J. L., Grytnes, J.-A., Jiang, Y.-H., Jetz, W., Kathleen Lyons, S., McCain, C. M., Magurran, A. E., Rahbek, C., Rangel, T. F. L. V. B., Soberón, J., Webb, C. O. and Willig, M. R., 2009. Patterns and causes of species richness: a general simulation model for macroecology. *Ecology Letters*, *12* (9), 873–886.

Jenkins, C. N. and Joppa, L., 2009. Expansion of the global terrestrial protected area system. *Biological Conservation*, *142* (10), 2166–2174.

Joppa, L. N. and Pfaff, A., 2009. High and Far: Biases in the Location of Protected Areas. *PLoS ONE*, *4* (12), e8273.

Keith, S. a, Newton, A. C., Morecroft, M. D., Bealey, C. E. and Bullock, J. M., 2009. Taxonomic homogenization of woodland plant communities over 70 years. *Proceedings. Biological Sciences / The Royal Society*, *276* (1672), 3539–3544.

Koleff, P., Gaston, K. J. and Lennon, J. J., 2003. Measuring beta diversity for presence-absence data. *Journal of Animal Ecology*, *72* (3), 367–382.

Laurance, W. F., Useche, D. C., Rendeiro, J., Kalka, M., Bradshaw, C. J. a, Sloan, S. P., Laurance, S. G., Campbell, M., Abernethy, K., Alvarez, P., Arroyo-Rodriguez, V., Ashton, P., Benítez-Malvido, J., Blom, A., Bobo, K. S., Cannon, C. H., Cao, M., Carroll, R., Chapman, C., Coates, R., Cords, M., Danielsen, F., De Dijn, B., Dinerstein, E., Donnelly, M. a, Edwards, D., Edwards, F., Farwig, N., Fashing, P., Forget, P.-M., Foster, M., Gale, G., Harris, D., Harrison, R., Hart, J., Karpanty, S., Kress, W. J., Krishnaswamy, J., Logsdon, W., Lovett, J., Magnusson, W., Maisels, F., Marshall, A. R., McClearn, D., Mudappa, D., Nielsen, M. R., Pearson, R., Pitman, N., van der Ploeg, J., Plumptre, A., Poulsen, J., Quesada, M., Rainey, H., Robinson, D., Roetgers, C., Rovero, F., Scatena, F., Schulze, C., Sheil, D., Struhsaker, T., Terborgh, J., Thomas, D., Timm, R., Urbina-Cardona, J. N., Vasudevan, K., Wright, S. J., Arias-G, J. C., Arroyo, L., Ashton, M., Auzel, P., Babaasa, D., Babweteera, F., Baker, P., Banki, O., Bass, M., Bila-Isia, I., Blake, S., Brockelman, W., Brokaw, N., Brühl, C. a, Bunyavejchewin, S., Chao, J.-T., Chave, J., Chellam, R., Clark, C. J., Clavijo, J., Congdon, R., Corlett, R., Dattaraja, H. S., Dave, C., Davies, G., Beisiegel, B. D. M., da Silva, R. D. N. P., Di Fiore, A., Diesmos, A., Dirzo, R., Doran-Sheehy, D., Eaton, M., Emmons, L., Estrada, A., Ewango, C., Fedigan, L., Feer, F., Fruth, B., Willis, J. G., Goodale, U., Goodman, S., Guix, J. C., Guthiga, P., Haber, W., Hamer, K., Herbinger, I., Hill, J., Huang, Z., Sun, I. F., Ickes, K., Itoh, A., Ivanauskas, N., Jackes, B., Janovec, J., Janzen, D., Jiangming, M., Jin, C., Jones, T., Justiniano, H., Kalko, E., Kasangaki, A., Killeen, T., King, H., Klop, E., Knott, C., Koné, I., Kudavidanage, E., Ribeiro, J. L. D. S., Lattke, J., Laval, R., Lawton, R., Leal, M., Leighton, M., Lentino, M., Leonel, C., Lindsell, J., Ling-Ling, L., Linsenmair, K. E., Losos, E., Lugo, A., Lwanga, J., Mack, A. L., Martins, M., McGraw, W. S., McNab, R., Montag, L., Thompson, J. M., Nabe-Nielsen, J., Nakagawa, M., Nepal, S., Norconk, M., Novotny, V., O'Donnell, S., Opiang, M., Ouboter, P., Parker, K., Parthasarathy, N., Pisciotta, K., Prawiradilaga, D., Pringle, C., Rajathurai, S., Reichard, U., Reinartz, G., Renton, K., Reynolds, G., Reynolds, V., Riley, E., Rödel, M.-O., Rothman, J., Round, P., Sakai, S., Sanaiotti, T., Savini, T., Schaab, G., Seidensticker, J., Siaka, A., Silman, M. R., Smith, T. B., de Almeida, S. S., Sodhi, N., Stanford, C., Stewart, K., Stokes, E., Stoner, K. E., Sukumar, R., Surbeck, M., Tobler, M., Tscharntke, T., Turkalo, A., Umapathy, G., van Weerd, M., Rivera, J. V., Venkataraman, M., Venn, L., Verea, C., de Castilho, C. V., Waltert, M., Wang, B., Watts, D., Weber, W., West, P., Whitacre, D., Whitney, K., Wilkie, D., Williams, S., Wright, D. D., Wright, P., Xiankai, L., Yonzon, P. and Zamzani, F., 2012. Averting biodiversity collapse in tropical forest protected areas. *Nature*, *489* (7415), 290–294.

Lennon, J. J., Beale, C. M., Reid, C. L., Kent, M. and Pakeman, R. J., 2011. Are richness patterns of common and rare species equally well explained by environmental variables? *Ecography*, *34* (4), 529–539.

Lewis, O. T., 2009. Biodiversity change and ecosystem function in tropical forests. *Basic and Applied Ecology*, *10* (2), 97–102.

Lima, M. M. and Mariano-Neto, E., 2014. Extinction thresholds for Sapotaceae due to forest cover in Atlantic Forest landscapes. *Forest Ecology and Management*, *312* , 260–270.

Loiselle, B. a., Jørgensen, P. M., Consiglio, T., Jiménez, I., Blake, J. G., Lohmann, L. G. and Montiel, O. M., 2007. Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcomes? *Journal of Biogeography*, *35*, 105-116.

McCafferty, W. P., 1998. Ephemeroptera and the Great American Interchange. *Journal of the North American Benthological Society*, *17* (1), 1–20.

McKnight, M. W., White, P. S., McDonald, R. I., Lamoreux, J. F., Sechrest, W., Ridgely, R. S. and Stuart, S. N., 2007. Putting beta-diversity on the map: broad-scale congruence and coincidence in the extremes. *PLoS Biology*, *5* (10), e272.

Naughton-Treves, L., Holland, M. B. and Brandon, K., 2005. the Role of Protected Areas in Conserving Biodiversity and Sustaining Local Livelihoods. *Annual Review of Environment and Resources*, *30* (1), 219–252.

Olden, J. D., 2006. Biotic homogenization: a new research agenda for conservation biogeography. *Journal of Biogeography*, *33* (12), 2027–2039.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'amico, J. a., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P. and Kassem, K. R., 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth. *BioScience*, *51* (11), 933-938.

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., Raven, P. H., Roberts, C. M. and Sexton, J. O., 2014. The biodiversity of species and their rates of extinction, distribution, and protection. *Science (New York, N.Y.)*, *344* (6187), 987-997.

Pimm, S. L. and Raven, P., 2000. Extinction by numbers. *Nature*, *403* , 843–845.

Pitman, N. C. A., Terborgh, J. W., Silman, M. R., V, P. N., A, D., Cerón, C. E., Palacios, W. A. and Aulestia, M., 2001. Dominance and Distribution of Tree Species in Upper Amazonian Terra Firme Forests. *Ecology*, *82* (8), 2101–2117.

Pressey, R. L., Cabeza, M., Watts, M. E., Cowling, R. M. and Wilson, K. A., 2007. Conservation planning in a changing world. *Trends in Ecology & Evolution*, *22* (11), 583–592.

Redford, K. H., 1992. The Empty Forest. *BioScience*, *42* (6), 412–422.

Rosauer, D. F., Ferrier, S., Williams, K. J., Manion, G., Keogh, J. S. and Laffan, S. W., 2014. Phylogenetic generalised dissimilarity modelling: a new approach to analysing and predicting spatial turnover in the phylogenetic composition of communities. *Ecography*, *37* , 21–32.

Soutullo, A., De Castro, M. and Urios, V., 2007. Linking political and scientifically derived targets for global biodiversity conservation: implications for the expansion of the global network of protected areas. *Diversity and Distributions*, *14* (4), 604–613.

Svenning, J.-C., Fløjgaard, C. and Baselga, A., 2011. Climate, history and neutrality as drivers of mammal beta diversity in Europe: insights from multiscale deconstruction. *The Journal of Animal Ecology*, *80 (2)*, 393-402.

Ter Steege, H., Pitman, N. C. a, Sabatier, D., Baraloto, C., Salomão, R. P., Guevara, J. E., Phillips, O. L., Castilho, C. V, Magnusson, W. E., Molino, J.-F., Monteagudo, A., Núñez Vargas, P., Montero, J. C.,

Feldpausch, T. R., Coronado, E. N. H., Killeen, T. J., Mostacedo, B., Vasquez, R., Assis, R. L., Terborgh, J., Wittmann, F., Andrade, A., Laurance, W. F., Laurance, S. G. W., Marimon, B. S., Marimon, B.-H., Guimarães Vieira, I. C., Amaral, I. L., Brienen, R., Castellanos, H., Cárdenas López, D., Duivenvoorden, J. F., Mogollón, H. F., Matos, F. D. D. A., Dávila, N., García-Villacorta, R., Stevenson Diaz, P. R., Costa, F., Emilio, T., Levis, C., Schietti, J., Souza, P., Alonso, A., Dallmeier, F., Montoya, A. J. D., Fernandez Piedade, M. T., Araujo-Murakami, A., Arroyo, L., Gribel, R., Fine, P. V. A., Peres, C. A., Toledo, M., Aymard C, G. A., Baker, T. R., Cerón, C., Engel, J., Henkel, T. W., Maas, P., Petronelli, P., Stropp, J., Zartman, C. E., Daly, D., Neill, D., Silveira, M., Paredes, M. R., Chave, J., Lima Filho, D. D. A., Jørgensen, P. M., Fuentes, A., Schöngart, J., Cornejo Valverde, F., Di Fiore, A., Jimenez, E. M., Peñuela Mora, M. C., Phillips, J. F., Rivas, G., van Andel, T. R., von Hildebrand, P., Hoffman, B., Zent, E. L., Malhi, Y., Prieto, A., Rudas, A., Ruschell, A. R., Silva, N., Vos, V., Zent, S., Oliveira, A. A., Schutz, A. C., Gonzales, T., Trindade Nascimento, M., Ramirez-Angulo, H., Sierra, R., Tirado, M., Umaña Medina, M. N., van der Heijden, G., Vela, C. I. A., Vilanova Torre, E., Vriesendorp, C., Wang, O., Young, K. R., Baider, C., Balslev, H., Ferreira, C., Mesones, I., Torres-Lezama, A., Urrego Giraldo, L. E., Zagt, R., Alexiades, M. N., Hernandez, L., Huamantupa-Chuquimaco, I., Milliken, W., Palacios Cuenca, W., Pauletto, D., Valderrama Sandoval, E., Valenzuela Gamarra, L., Dexter, K. G., Feeley, K., Lopez-Gonzalez, G. and Silman, M. R., 2013. Hyperdominance in the Amazonian tree flora. *Science (New York, N.Y.)*, *342*, 325-334.

Thomas, C. D. and Gillingham, P. K., 2015. The performance of p rotected a reas for biodiversity under climate change., 718–730.

Tobler, M., Honorio, E., Janovec, J. and Reynel, C., 2007. Implications of collection patterns of botanical specimens on their usefulness for conservation planning: an example of two neotropical plant families (Moraceae and Myristicaceae) in Peru. *Biodiversity and Conservation*, *16* (3), 659–677.

Vermeij, G. J., 1991. When Biotas Meet: Understanding Biotic Interchange. *Science*, *253* , 1099–1104.

Wilkie, D. S., Bennett, E. L., Peres, C. A. and Cunningham, A. A., 2011. The empty forest revisited. *Annals of the New York Academy of Sciences*, *1223* , 120–128.

7. APPENDICES

Appendix I. Summary tables for the three data sub-samples used in **Chapter 2**.

Table 1. Species with highest occurrences in sample 1 data set. Count of occurrence refers to the number of times a species occurs in a different grid cell. It can be thougth of as the species range of occurrence within the sample.

Species	**Count**
Hamelia patens	183
Casearia corymbosa	161
Guazuma ulmifolia	153
Lantana camara	149
Dendropanax arboreus	145
Bursera simaruba	143
Malvaviscus arboreus	135
Byrsonima crassifolia	131
Cissampelos pareira	125
Trema micrantha	117
Psychotria pubescens	116
Iresine diffusa	112
Conostegia xalapensis	110
Inga vera	108
Solanum aphyodendron	108
Aphelandra scabra	107
Acacia angustissima	105
Neurolaena lobata	105
Piper obliquum	103
Acalypha diversifolia	101
Chiococca alba	101
Chromolaena odorata	101
Vernonanthura patens	101
Singletons	166
Doubletons	172
Average	17
No. species with count of occurrence >10	1190
No. species with count of occurrence ≤10	1226
Total	2416

Table 2. Species with highest occurrences in sample 2 data set. Count of occurrence refers to the number of times a species occurs in a different grid cell. It can be thougth of as the species range of occurrence within the sample.

Species	**Count**
Hamelia patens	195
Guazuma ulmifolia	165
Casearia corymbosa	160
Lantana camara	158
Dendropanax arboreus	148
Malvaviscus arboreus	143
Inga vera	140
Byrsonima crassifolia	138
Bursera simaruba	136
Trema micrantha	129

Aphelandra scabra	126
Psychotria pubescens	121
Conostegia xalapensis	119
Cissampelos pareira	117
Solanum aphyodendron	116
Vernonanthura patens	115
Ardisia compressa	114
Chromolaena odorata	114
Iresine diffusa	107
Guarea glabra	105
Acacia angustissima	103
Singletons	180
Doubletons	153
Average	17
No. species with count of occurrence >10	1128
No. species with count of occurrence <11	1289
Total	2417

Table 3. Species with highest occurrences in sample 3 data set. Count of occurrence refers to the number of times a species occurs in a different grid cell. It can be thougth of as the species range of occurrence within the sample.

Species	**Count**
Hamelia patens	202
Casearia corymbosa	188
Lantana camara	166
Bursera simaruba	162
Guazuma ulmifolia	158
Malvaviscus arboreus	146
Psychotria pubescens	143
Byrsonima crassifolia	138
Chiococca alba	136
Inga vera	136
Trema micrantha	135
Iresine diffusa	134
Aphelandra scabra	132
Dendropanax arboreus	128
Conostegia xalapensis	126
Solanum aphyodendron	125
Cissampelos pareira	123
Chromolaena odorata	119
Solanum americanum	116
Piper obliquum	110
Acacia angustissima	109
Muntingia calabura	107
Rollinia mucosa	104
Vernonanthura patens	103
Cissus microcarpa	100
Guarea glabra	100
Trichostigma octandrum	100
Singletons	153
Doubletons	152
Average	18
No. species with count of occurrence >10	1225
No. species with count of occurrence <11	1185
Total	2410

Appendix II. Resulting % DE and Sum of coefficients for all model combinations implemented in **Chapter 2**.

Model	% Deviance Explained	Sum of Coefficients (amount of predicted turnover attributable to this predictor)
Sorensen		
Geographic Distance	11.993097	2.892450
Temperature (Bio 1)	7.066383	2.470855
DEM	7.840649	2.576098
Precipitation (Bio 12)	7.247456	2.934669
Geo dist + Precipitation (Bio 12)	16.096856	
Geo dist		2.128
Bio 12		2.364
Geo dist + Temp (Bio 1)	17.070561	
Geo dist		2.513
Bio 1		2.083
Geo dist + DEM	17.493475	
Geo dist		2.522
DEM		2.333
Temp (Bio 1) + Precipitation (Bio 12)	13.687857	
Temp (Bio 1)		2.411
Precipitation (Bio 12)		2.624
Temp + DEM	8.363417	
Temp		0.986
DEM		2.323
DEM + Precipitation (Bio 12)	14.225177	
DEM		2.453
Precipitation		2.579
Geo dist + DEM + Temp	17.722766	
Geo dist		2.46

DEM		1.954
Temp		0.801
Geo dist + DEM + Precipitation	21.036163	
Geo dist		1.837
DEM		2.231
Precipitation		2.077
Temp + Precipitation + DEM	14.752946	
Temp		1.122
Precipitation		2.56
DEM		2.008
Geo dist + Temp + Precipitation	20.774690	
Geo dist		1.783
Temp		2.012
Precipitation		2.13
All 4 variables	21.312692	
Geo dist		1.764
Temp		0.945
DEM		1.693
Precipitation		2.082
All 19 bioclim variables + Geo Distance + DEM The predictors which most contributed to the model were, in order of contribution:	22.78	
Geo distance		1.447
DEM		1.362
bio 5=max temp/warmest month		1.344
bio 16=precipn/wettest quarter		
bio 15=precipn seasonality		
bio 8=mean temp/wettest quarter		
bio 12=annual preciptn		
bio 19=precipn/coldest quarter		
Simpson Beta		
Geodist	6.189247	2.480506

Temp (Bio 1)	5.004102	2.502871
DEM	6.122225	2.562301
Precipitation (Bio2)	4.607019	2.628039
Geodist+DEM	10.912250	
Geodist		2.1290740
DEM		2.5405770
Geodist + Precip	9.113565	
Geodist		1.8376660
Precip		2.2200250
Geodist+Temp	10.078491	
Geodist		2.1589
Temp		2.3805
DEM+Precip	9.789758	
DEM		2.3299690
Precip		2.1910490
DEM+Temp	6.596633	
DEM		2.5571
Temp		0.8506
Temp+Precip	8.848338	
Temp		2.2077
Precip		2.249
Temp+Precip+DEM	10.180541	
Temp		0.790455
Precip		2.12108
DEM		2.3388
Geodist+DEM+Precipitation	13.102301	
Geodist		1.5450190
DEM		2.2881680
Precip		1.8456430
Geodist+Temp+Precip	12.414317	

Geodist		1.5451010
Temp		2.1094500
Precip		1.9101840
Geodist+DEM+Temp	11.166027	
Geodist		2.0540000
DEM		2.5326660
Temp		0.631571
All 4 variables	13.331448	
Geo dist		1.4761
Temp		0.6115
DEM		2.288
Precipitation		1.808
All 19 bioclim variables + Geo Distance + DEM The predictors which most contributed to the model were, in order of contribution:	15.029840	
Geo distance		1.048495
DEM		1.949695
Bio 16 precip/wettest quarter		0.864368
Bio 5 max temp/warmest month		0.839051
Bio 15 precip seasonality		0.827619
Bio 12 annual precipitation		0.621005
Bio 8		0.493776
Bio 2b		0.43245
Bio 19		0.282526
Bio 9		0.161305

Appendix III. Proportion of plant habitat protected after accounting for habitat loss to 2000. (**Chapter 4**)

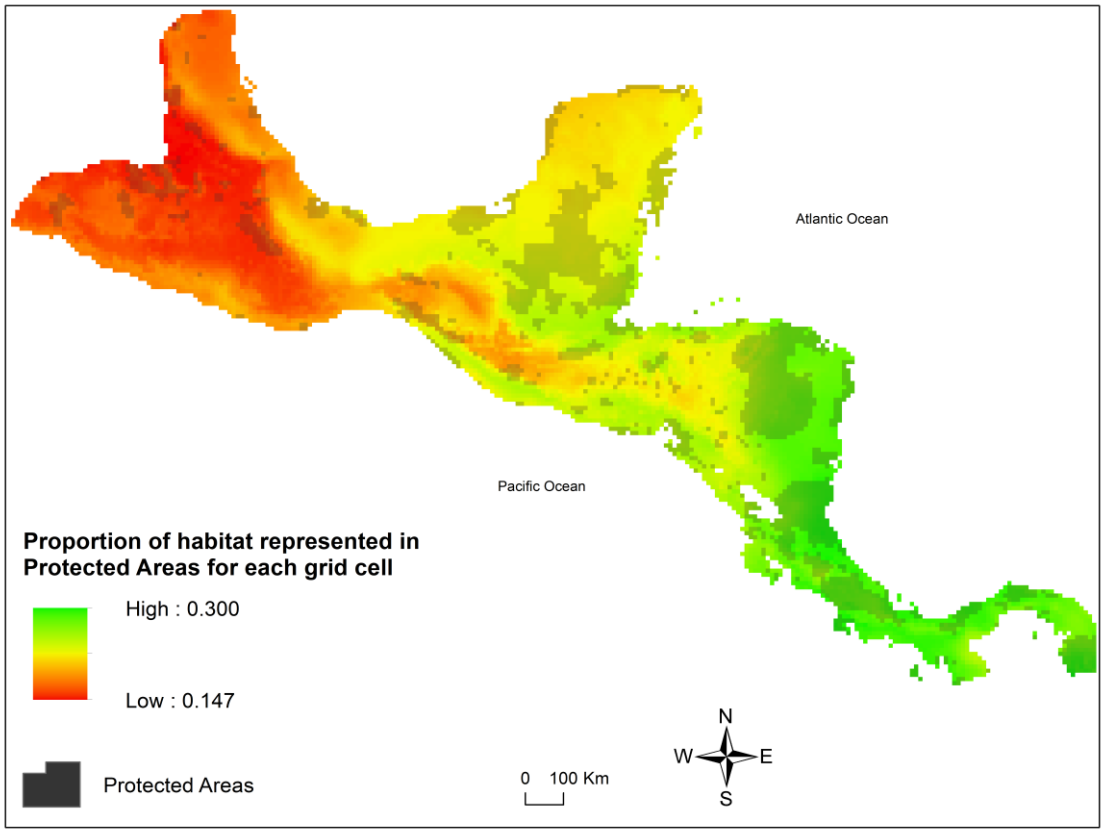


Figure 1. Proportion of vascular plant habitat in protected areas across Mesoamerica for the four predictor model, after habitat loss to 2000. Green indicates that a high proportion of areas predicted to be similar to that site are within reserves. Red indicates poor representation of similar areas within reserves.

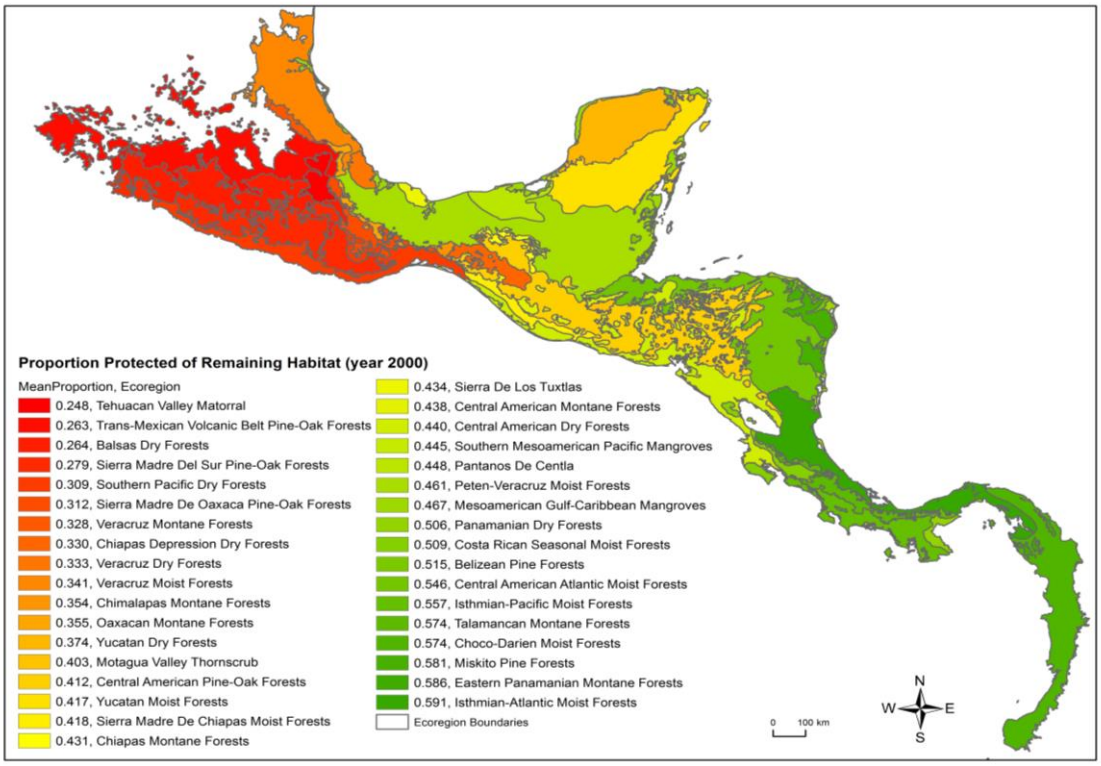
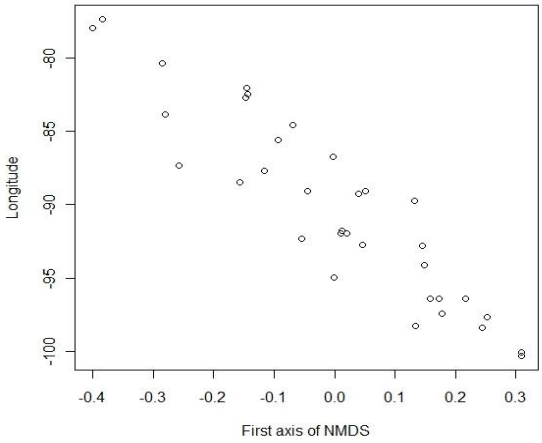


Figure 2. Mean proportion of habitat represented in current protected area network for each ecoregion, after habitat loss up to year 2000 is accounted for. Ecoregions are sorted in ascending order of habitat representation.

Appendix IV. Summary of data stratified by the 34 WWF ecoregions used in **Chapter 5** analysis.

Ecoregion	No. of records	No. of species	No. of genera	No. of families	Area of ecoregion (km^2)
Isthmian-Atlantic Moist Forests	76503	1691	648	142	na
Isthmian-Pacific Moist Forests	50812	1543	632	140	28881
Talamancan Montane Forests	40731	1384	561	137	15991
Peten-Veracruz Moist Forests	38878	1461	588	135	149100
Central American Dry Forests	30547	1100	501	124	50101
Costa Rican Seasonal Moist Forests	28694	1307	571	141	10620
Central American Pine-Oak Forests	22845	1342	553	134	127910
Yucatan Moist Forests	17072	609	334	88	64,012
Trans-Mexican Volcanic Belt Pine-Oak Forests	16456	687	354	103	72802
Southern Pacific Dry Forests	16306	900	443	115	16400
Veracruz Moist Forests	15921	836	423	123	69100
Central American Atlantic Moist Forests	14553	1183	528	132	155020
Sierra De Los Tuxtlas	12133	798	416	116	na
Yucatan Dry Forests	11951	497	282	75	45554
Oaxacan Montane Forests	10301	713	348	116	na
Choco-Darien Moist Forests	9648	1039	470	108	82079
Mesoamerican Gulf-Caribbean Mangroves	9631	1027	483	111	na
Veracruz Dry Forests	9288	575	328	90	35546
Balsas Dry Forests	7947	640	330	100	161098
Central American Montane Forests	6156	837	403	128	7676
Sierra Madre Del Sur Pine-Oak Forests	5878	741	373	111	41129
Sierra Madre De Oaxaca Pine-Oak Forests	3374	593	330	105	na
Sierra Madre De Chiapas Moist Forests	3331	667	360	112	na
Chiapas Depression Dry Forests	3006	637	334	90	na
Pantanos De Centla	2812	368	235	73	
Veracruz Montane Forests	2120	399	247	89	6615
Southern Mesoamerican Pacific Mangroves	2032	598	367	99	na
Eastern Panamanian Montane Forests	1982	554	297	95	2905
Chiapas Montane Forests	1953	498	270	95	na
Panamanian Dry Forests	1952	470	290	89	5010
Belizian Pine Forests	1332	404	248	83	2750
Miskito Pine Forests	1050	310	196	79	15064
Chimalapas Montane Forests	875	319	193	81	na
Motagua Valley Thornscrub	704	193	138	57	2363

Appendix V. Spearman's Rank Correlation Tests **(Chapter 5)**



Spearman's rank correlation rho

data: ord$points[, 1] and climate$Long

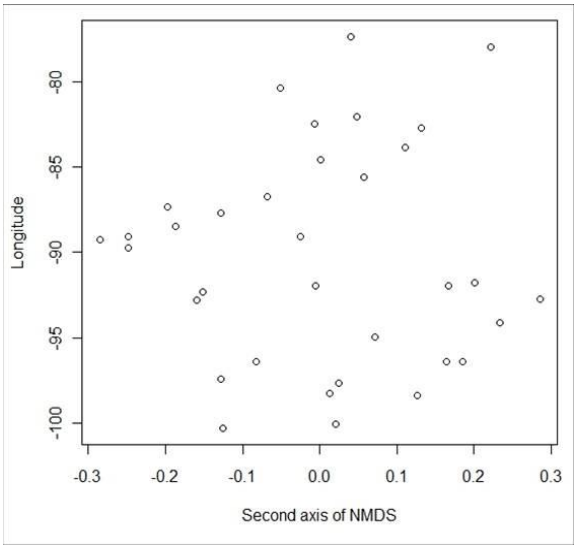
S = 12560, p-value < 2.2e-16

alternative hypothesis: true rho is not equal to 0

sample estimates:

rho

-0.9190222



Spearman's rank correlation rho

data: ord$points[, 2] and climate$Long

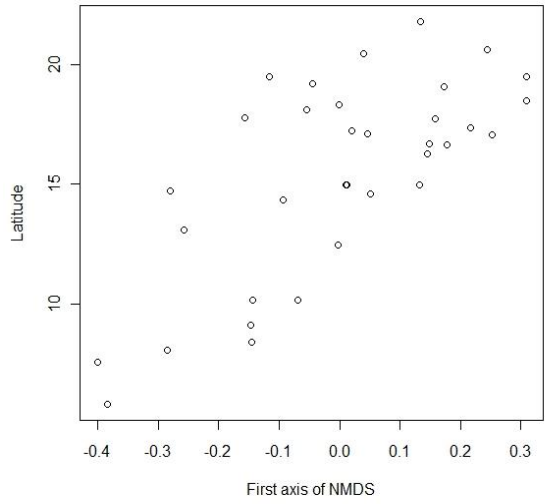
S = 6748, p-value = 0.8617

alternative hypothesis: true rho is not equal to 0

sample estimates:

rho

-0.03101604



Spearman's rank correlation rho

data: ord$points[, 1] and climate$Lat

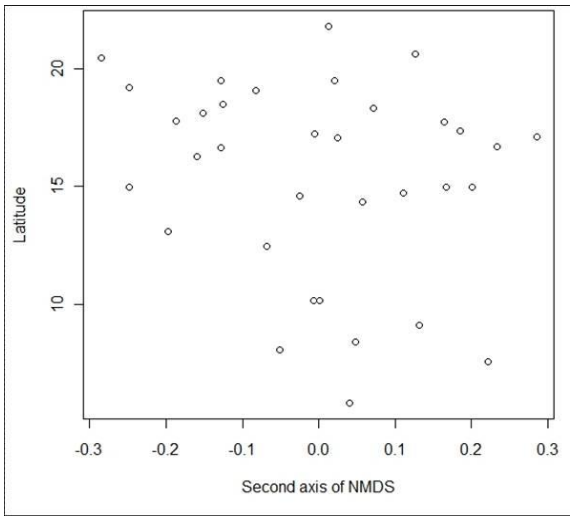
S = 2446, p-value = 0.0001052

alternative hypothesis: true rho is not equal to 0

sample estimates:

rho

0.6262796



Spearman's rank correlation rho

data: ord$points[, 2] and climate$Lat

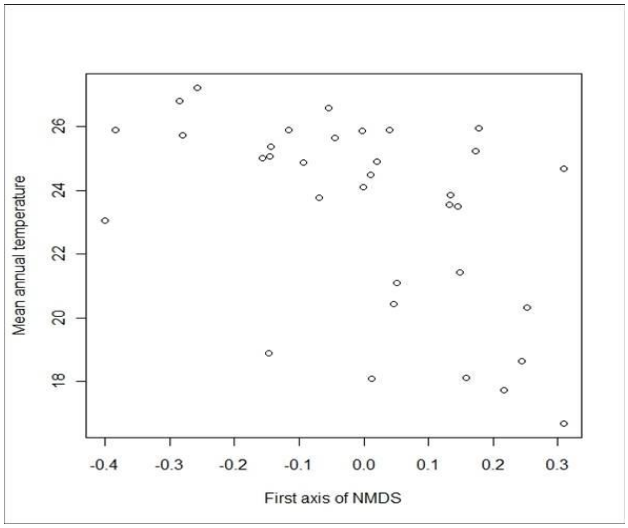
S = 8034, p-value = 0.195

alternative hypothesis: true rho is not equal to 0

sample estimates:

rho

-0.2275019



Spearman's rank correlation rho

data: ord$points[, 1] and climate$meant

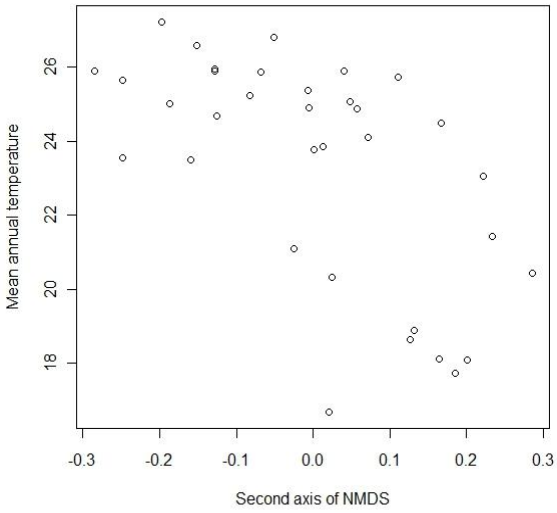
S = 9896, p-value = 0.002245

alternative hypothesis: true rho is not equal to 0

sample estimates:

rho

-0.5119939



Spearman's rank correlation rho

data: ord$points[, 2] and climate$meant

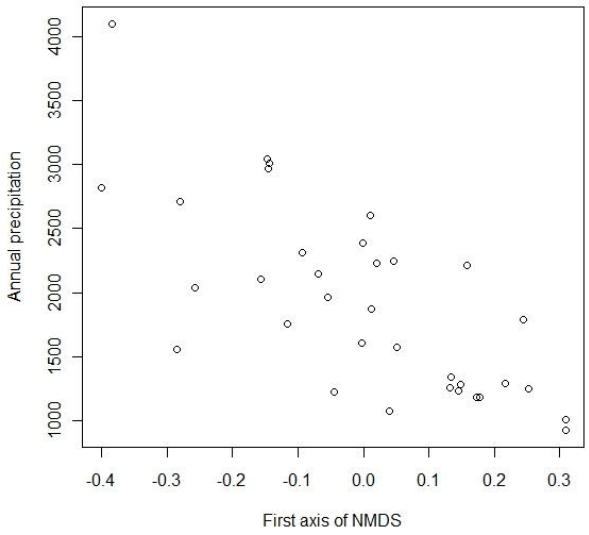
S = 10628, p-value = 0.0001137

alternative hypothesis: true rho is not equal to 0

sample estimates:

rho

-0.623835



Spearman's rank correlation rho

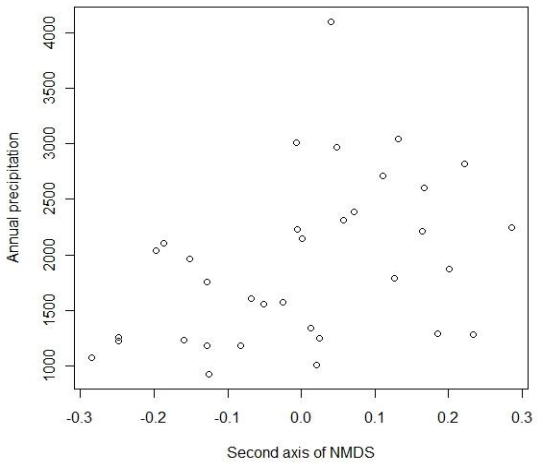
data: ord$points[, 1] and climate$anprec

S = 11144, p-value = 6.965e-06

alternative hypothesis: true rho is not equal to 0

sample estimates:

rho -0.7026738



Spearman's rank correlation rho

data: ord$points[, 2] and climate$anprec

S = 3322, p-value = 0.003448

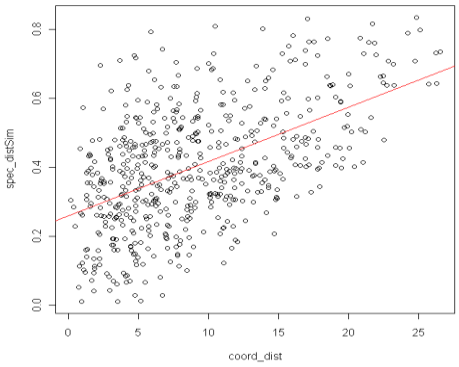
alternative hypothesis: true rho is not equal to 0

sample estimates:

rho

0.492437

MANTEL TEST RESULTS (Chapter 5)



intercept=0.2582, intercept=0.0158

Call:

lm(formula = spec_distSim ~ coord_dist)

Residuals:

Min	1Q	Median	3Q	Max
-0.33566	-0.11556	-0.01449	0.10407	0.44157

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.258160	0.011642	22.18	<2e-16 ***
coord_dist	0.015823	0.001082	14.62	**<2e-16 *****

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1481 on 559 degrees of freedom

Multiple R-squared: 0.2767, Adjusted R-squared: 0.2754

F-statistic: 213.9 on 1 and 559 DF, p-value: < 2.2e-16

Distance in Species composition vs distance in Euclidean distance (geographic space):

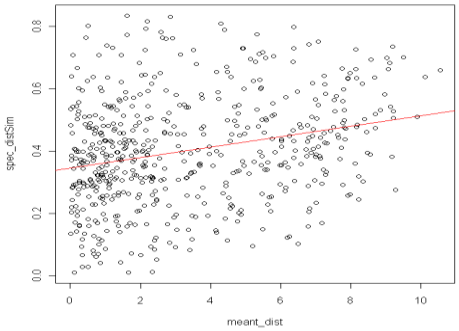
Mantel statistic based on Spearman's rank correlation rho

Call:

mantel(xdis = spec_distSim, ydis = coord_dist, method = "spearman")

Mantel statistic r: 0.4831

Significance: 0.001



intercept=0.3454, slope=0.0169

Call:

lm(formula = spec_distSim ~ meant_dist)

Residuals:

Min	1Q	Median	3Q	Max
-0.37222	-0.10564	-0.00527	0.11269	0.46295

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.345407	0.011313	30.533	< 2e-16 ***
meant_dist	0.016897	0.002641	6.399	**3.31e-10 *****

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1681 on 559 degrees of freedom

Multiple R-squared: 0.06825, Adjusted R-squared: 0.06659

F-statistic: 40.95 on 1 and 559 DF, p-value: 3.309e-10

Distance in Species composition vs distance in Mean Temp:

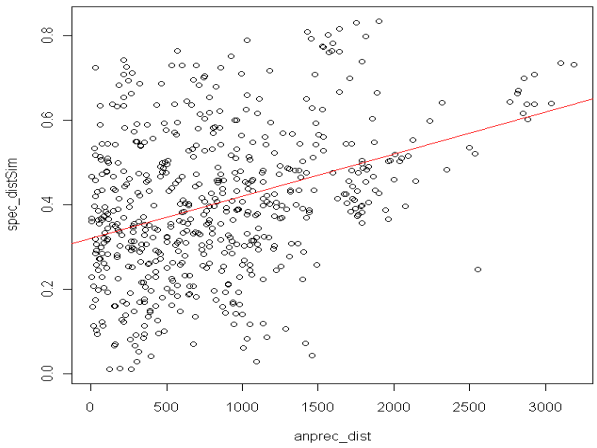
Mantel statistic based on Spearman's rank correlation rho

Call:

mantel(xdis = spec_distSim, ydis = meant_dist, method = "spearman")

Mantel statistic r: 0.2544

Significance: 0.001



intercept=0.3202, slope=0.0001

Call:

lm(formula = spec_distSim ~ anprec_dist)

Residuals:

Min	1Q	Median	3Q	Max
-0.42114	-0.10576	-0.01319	0.10022	0.40252

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.202e-01	1.124e-02	28.495	<2e-16 ***
anprec_dist	9.842e-05	1.074e-05	9.166	**<2e-16 *****

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1624 on 559 degrees of freedom

Multiple R-squared: 0.1307, Adjusted R-squared: 0.1291

F-statistic: 84.01 on 1 and 559 DF, p-value: < 2.2e-16

Distance in Species composition vs distance in Annual Precipitation:

Mantel statistic based on Spearman's rank correlation rho

Call:

mantel(xdis = spec_distSim, ydis = anprec_dist, method = "spearman")

Mantel statistic r: 0.3222

Significance: 0.001

Partial Mantel statistic: Species composition vs Mean Temp accounting for geographic distance

Call:

mantel.partial(xdis = spec_distSim, ydis = meant_dist, zdis = coord_dist, method = "spearman")

Mantel statistic r: 0.2795

Significance: 0.001

Partial Mantel statistic Species composition vs Annual Precipitation accounting for geographic distance

Call:

mantel.partial(xdis = spec_distSim, ydis = anprec_dist, zdis = coord_dist, method = "spearman")

Mantel statistic r: 0.1708

Significance: 0.003